

Improved Event Location Uncertainty Estimates

**István Bondár
Keith McLaughlin
Hans Israelsson**

**Science Applications International Corporation
10260 Campus Drive
San Diego, CA, 92121**

Final Report

30 June 2008

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.



**AIR FORCE RESEARCH LABORATORY
Space Vehicles Directorate
29 Randolph Road
AIR FORCE MATERIEL COMMAND
Hanscom AFB, MA 01731-3010**

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release and is available to the general public, including foreign nationals. Qualified requestors may obtain additional copies from the Defense Technical Information Center (DTIC) (<http://www.dtic.mil>). All others should apply to the National Technical Information Service.

AFRL-RV-HA-TR-2008-1074 HAS BEEN REVIEWED AND IS APPROVED FOR
PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//Signature//

ROBERT RAISTRICK
Contract Manager

//Signature//

PAUL TRACY, Acting Chief
Battlespace Surveillance Innovation Center

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 30-06-2008		2. REPORT TYPE Final Report		3. DATES COVERED (From - To) 04-08-2005 to 11-07-2008	
4. TITLE AND SUBTITLE Improved Event Location Uncertainty Estimates				5a. CONTRACT NUMBER FA8718-05-C-0018	
				5b. GRANT NUMBER N/A	
				5c. PROGRAM ELEMENT NUMBER 62601F	
6. AUTHOR(S) I. Bondár, K. McLaughlin and H. Israelsson				5d. PROJECT NUMBER 1010	
				5e. TASK NUMBER SM	
				5f. WORK UNIT NUMBER A1	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Science Applications International Corporation 10260 Campus Drive San Diego, CA 92121				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory 29 Randolph Rd. Hanscom AFB, MA 01731-3010				10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/RVBYE	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-RV-HA-TR-2008-1074	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The objective of this project was to develop methodologies that improve location uncertainties in the presence of correlated, systematic model errors and non-Gaussian measurement errors. We have developed a methodology based on copula theory to obtain robust estimates of variogram models for travel-time error. Using this methodology, we produced generic, transportable variogram models for regional Pn and teleseismic P phases. These models are employed to estimate the network covariance matrix that describes the spatial correlation structure of travel-time predictions errors. We have developed and validated models of measurement errors that map phase picking delay and variance as a function of SNR, as well as a function of epicentral distance and body wave magnitude considered as a surrogate for SNR. The phase pick delays are implemented as travel-time corrections, while the measurement error variances add to the diagonal of the network covariance matrix to obtain the full, non-diagonal data covariance matrix. Our representation of model and measurement errors assumes that the bulk of 3D velocity heterogeneities in the Earth are accounted for by calibrated travel-time predictions. We have developed a linearized iterative location algorithm that makes use of the full data covariance matrix, hence accounts for the correlated model error structure. By relocating a large number of GT events (GT0-2 nuclear explosions, and GT5 earthquakes produced by Reciprocal Cluster Analysis we show that ignoring the correlated error structure leads to rapidly deteriorating error ellipse coverage and location bias with increasingly correlated networks. Calibrated travel-times are responsible for the bulk of location improvements, but they do little to improve error ellipse coverage. While the effect of measurement errors on event locations is small on the order of 1-2 km), they tend to tighten event clusters by reducing the systematic separation between large and small events. We demonstrate that taking into account the correlated model error structure significantly improves error ellipse coverage, and for unbalanced networks, reduces location bias. We also show that the deteriorating effect of non-Gaussian error distributions, albeit consistent and non-negligible, is of secondary importance compared to the penalty paid for ignoring the correlation structure in travel-time predictions.					
15. SUBJECT TERMS Correlated model errors, Location method					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 82	19a. NAME OF RESPONSIBLE PERSON Robert Raistrick
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (include area code) 781-377-3726

Contents

1. Introduction	1
2. Background	2
3. Technical Approach	6
3.1. Location algorithm with correlated errors	6
3.1.1. Model errors: Estimation of the network covariance matrix	6
3.1.2. Measurement errors: Bias and variance	10
3.1.2.1. NTS data	10
3.1.2.2. Measurement error model based on amplitude/period ratio	11
3.1.2.3. Measurement error model based on mb and epicentral distance	14
3.1.2.4. Measurement error model based on SNR	20
3.1.3. Location algorithm	25
3.1.3.1. Baseline location algorithm	25
3.1.3.2. Improved location algorithm using full data covariance matrix	26
4. Validation tests	29
4.1. Measurement errors	30
4.1.1. mb-distance dependent bias and variance	30
4.1.1.1. Relocation of 279 Yucca Flat explosions	30
4.1.1.2. mb dependence for other test sites	31
4.1.1.3. Transportability of the NTS mb-based measurement error model	33
4.1.2. SNR-dependent bias and variance	37
4.2. Correlated model errors	37
4.2.1. Coverage with increasing number of correlated stations	37

4.2.2. Relocation of a large set of GT5 earthquakes	38
4.2.3. Sparse, unbalanced networks	42
4.3. Correlated model error structure vs non-Gaussian errors	44
4.3.1. Dense networks	46
4.3.2. Increasingly correlated networks	50
4.4. Correlated model errors with mb-based measurement errors and phase pick delay correction	53
4.4.1. Nuclear explosions	54
4.4.2. Earthquakes	57
5. Discussion and conclusions	59
Acknowledgements	63
References	65
Appendix A	69
A.1 Generating correlated multivariate non-Gaussian random numbers	69
A.2 Robust variogram model estimation	70

Figures

1. a) Teleseismic stations (28° - 90°) that recorded the 7 October, 1994 underground nuclear explosion at the Lop Nor, China test site. b) Trajectory of median mislocation using subnetworks starting with 6-station networks and gradually increasing to 400 stations (solid triangle). As the number of stations in the subnetworks with optimal azimuthal coverage increases, the location is driven away from the ground truth (star) location.
2. a) Error ellipse coverage, b) error ellipse area and c) mislocation with increasing number of stations for the 7 October, 1994 Lop Nor explosion. Because of the correlated model error structure, the location moves away from the true location, the area of the error ellipse monotonically decreases and the error ellipse no longer covers the true location (coverage > 1) with increasing number of stations.
3. a) Pn GT0 residuals (observed - iasp91 predictions with respect to the GT0 locations and origin times) for explosions at Yucca Flat, Nevada Test Site plotted as a function of magnitude, mb. Circles represent the median residuals at individual stations. The thick line connects the median station residuals. The non-zero median indicates that the iasp91 velocity model is too fast for the Western US. b) Pn GT0 residuals from CUB2 model predictions. The CUB2 model reduces the path effects not modeled by iasp91 and reveals a trend (dashed line) due to systematically late picks with decreasing event size. The median bias decreases by 0.3 seconds from mb 4.5 to 5.5.
4. Globally distributed GT5 earthquake clusters (circles) and GT0-2 clusters of nuclear explosions (stars).
5. Variogram models for a) regional Pn and b) teleseismic P derived from earthquake (blue) and explosion (magenta) clusters. The thick red lines show the generic variogram models derived from all clusters.
6. a) Arrival time residuals as a function of amplitude/period ratios (A/T) for the station COL (College Outpost, Alaska) for NTS underground nuclear explosions. Different symbols are used for explosions at the two areas Pahute Mesa (open circles) and Yucca Flat (filled circles) of NTS. There is a systematic difference in arrival time residuals between the two areas as indicated by the estimated relations (solid and dashed lines) between arrival time residual as a function of A/T . b) The residuals for Pahute Mesa and Yucca Flat in Figure 2) have been combined after subtracting estimated path effects. The estimated bias and standard deviation in measurement error as a function of A/T are indicated as dashed and dotted lines.
7. Arrival time residuals as a function of amplitude/period (A/T) ratios at the two stations COL (College Outpost, Alaska) and EDM (Edmonton, Canada) for explosions at Yucca Flat. Dashed and solid lines mark the thresholds from Lilwall and Neary (1986) for COL and EDM respectively.

8. a) Arrival time residuals as a function of network mb, mb(ISC), for the station COL (College Outpost, Alaska) for NTS underground nuclear explosions. Different symbols are used for explosions at Pahute Mesa (open circles) and Yucca Flat (filled circles). There is a systematic difference in arrival time residuals between the two areas as indicated by the estimated relations (solid and dashed lines) between arrival time residual as a function of mb(ISC). b) Pahute Mesa and Yucca Flat data combined after normalization. The estimated bias and standard deviation in measurement error as a function of network mb are indicated as dashed and dotted lines.

9. Estimates of delay (BIAS) and standard deviation (SIGMA) in arrival times of NTS explosions a) as a function of network mb(ISC) for 60 stations, b) as a function of SNR relative to the Lilwall and Neary (1986) thresholds (SNR =1).

10. Histograms for arrival time residuals of Yucca Flat events compared with Gaussian distributions. a) iasp91 residuals. b) residuals with estimated path corrections. c) residuals with path corrections and mb-based delay corrections. c) normalized residuals: path and bias corrected residuals are divided by the mb-based standard deviation estimates. Note that the horizontal scale for the normalized residuals is in sample quantiles and the density function of the standard Gaussian (zero mean and unit standard deviation) is drawn as a dashed line.

11. The contours show the a) delay and b) standard deviation in arrival time measurements as a function of network mb and epicentral distance for NTS explosions.

12. a) Double difference (DD) residuals where all four SNR are similar. The running SMAD (solid curve) represents the estimate of total variance. b) Double difference residuals where the SNR of three arrivals is large and the SNR of one arrival is small. The running median (dashed line) provides an estimate of the reading error bias.

13. SNR dependent model for pick delay (dashed line) and standard deviation (circles) derived from GT5 earthquakes with arrival and SNR data reported in the IDC REB.

14. Histograms for arrival time residuals of GT5 events compared with Gaussian distributions. a) iasp91 residuals. b) residuals with SSSCs. c) residuals with SSSCs and SNR-based delay corrections. c) normalized residuals: path and bias corrected residuals are divided by the SNR-based standard deviation estimates. Note that the horizontal scale for the normalized residuals is in sample quantiles and the density function of the standard Gaussian (zero mean and unit standard deviation) is drawn as a dashed line.

15. Regional network (3° - 10°) for the 1992/03/26 Pahute Mesa, NTS explosion (star). Triangles denote stations with positive Pn residuals, inverted triangles show stations with negative Pn residuals. Despite the small azimuthal gap (60°), the network is heavily unbalanced, owing to dense local networks in the Los Angeles basin to the Southwest and around the Mt. St. Helens volcano to the Northwest. b) Median residual difference squares in 5-percentile bins calculated from the GT residuals of the 1992/03/25 Pahute Mesa event (light blue line). The blue line shows the variogram model derived from the

entire Pahute Mesa nuclear explosion cluster, the red line represents the generic Pn variogram model.

16. a) Station-station correlation matrix calculated from the regional network in Figure 5 and the Pn generic variogram model in Figure 4. The correlation matrix is arranged by the nearest-neighbor order of stations and exhibits a quasi block-diagonal structure. b) Cumulative eigenvalue spectrum of the network covariance matrix. The 40 largest eigenvalues explain 95% of the information carried by the 97-station network. c) Full data correlation matrix (with variances of reading errors added to the diagonal of the network covariance matrix). Reading errors weaken the correlation structure carried by the network. d) Because of the somewhat diluted correlation structure, more eigenvalues (83) are required to explain 95% of the covariance structure.

17. Relocation of the 1992/03/26 Pahute Mesa GT0 explosion (star) with iasp91 (blue) and CUB2 travel-time predictions assuming independent errors (green), and using CUB2 travel-times and the full data covariance matrix (red). When correlated errors are taken into account, the error ellipse covers the true location (star). The “eigenstations” represent a more balanced network resulting in a more circular error ellipse.

18. Differences in arrival time residuals for six stations (DOU, GRR, INK, KHC, NDI, WRA) relative to station HFS as a function of network mb(ISC). Estimated bias and standard deviations as a function of mb(ISC) are plotted as dashed and dotted lines.

19. a) Estimated delay and b) standard deviation as a function of network mb(ISC) for Degelen explosions at 41 stations.

20. Mislocations (km) of NTS (crosses), Balapan (inverted triangles), Lop Nor (circles) and North Caspian (Azgir, Vega; squares) underground nuclear explosions using Pn arrivals with a) uncalibrated travel-time predictions (iasp91); b) uncalibrated travel-time predictions and mb-based delay and variance estimates c) calibrated (CUB2) travel-time predictions; d) calibrated travel-time predictions and mb-based delay and variance estimates. The mb-based delay corrections and measurement error estimates tighten 5 out of 7 clusters as listed in Table 4.

21. Mislocations of the Izmit, Turkey earthquake cluster using a) uncalibrated (iasp91); b) calibrated (CUB2, J362D28) travel-time predictions; c) calibrated travel-time predictions and mb-based delay and variance estimates. d-f are the same as a-c but for the Kileaua, Hawaii earthquake cluster. Events are color-coded by their mb and locations are plotted relative to the GT5 locations in Easting-Northing coordinates (km). Error ellipses at the one and two sigma levels encompassing the point clouds are also shown. SSSCs are responsible for the bulk of location improvements; the mb-based delay corrections and weighting scheme further tighten the clusters.

22. Mislocation (bottom), error ellipse area (middle) and coverage (top) with increasing number of optimally distributed stations when correlated errors are ignored (blue lines) and when correlated errors are accounted for (red lines) for the a) 1994/10/07, Lop Nor,

and b) 1992/03/26, Pahute Mesa explosions. Coverage and error ellipse area are nearly impervious to the number of stations for the correlated case.

23. Cumulative distributions of GT5 earthquake a) coverage parameters, b) semi-major axes, and c) mislocations when correlated errors are ignored (blue) and when correlated errors are accounted for (red). Accounting for correlation increases actual coverage from about 68% to 85% and the cumulative more closely approaches the theoretical χ^2 distribution (black line in a).

24. Cumulative distributions of mislocations of GT5 earthquakes with iasp91 (green), with calibrated CUB2 and/or J362D28 travel-time predictions (blue), and with calibrated travel-times with correlated errors (red). The upper panel shows the events with $m_b \geq 4$. Calibrated travel-times are the first-order effects in location improvements; accounting for correlated model errors brings incremental location improvements.

25. a) 90% coverage, b) median error ellipse area, and c) median mislocation of 2275 GT5 earthquakes ordered by the number of stations used in the location when correlated errors are accounted for (red), and when they are ignored (blue). Coverage and error ellipse area are less dependent upon the number of stations for the correlated case.

26. Regional network (0° - 15°) for the Matochkin Shar, Novaya Zemlya test site (star). The numbers in the parentheses indicate the number of events (out of 29 GT1-2 explosions) reported by a particular station. Most of the stations are located in Fennoscandia, representing a sparse, heavily unbalanced network.

27. a) 90% error ellipse area, b) origin time difference, and c) mislocation for 29 GT1-2 Matochkin Shar nuclear explosions using iasp91 travel-times and assuming independent errors (green triangles), and using calibrated (CUB2) travel-times and accounting for the correlated model error structure (red circles). The horizontal bands represent the number of defining phases, indicated on the right hand axis.

28. Grid search misfit surfaces for the 1975/08/23 8:59:59 Novaya Zemlya underground nuclear explosion (red star) using regional Pn from KBS, KEV, KIR, SOD and KJF. Misfit contours from the grid search (thin lines) and 90% confidence error ellipses from the iterative least squared relocations (thick, colored lines and grey diamonds) are shown when a) uncalibrated (iasp91) travel-times assuming independent errors, b) calibrated (CUB2) travel-times assuming independent errors, and c) calibrated travel-times with correlated model error structure were used. The misfit surface is the steepest when correlated errors are accounted for.

29. GT residual distributions of a) regional Pn and b) teleseismic P phases from the GT5 validation data set. Solid and dashed lines represent the best fitting Generalized Extreme Value (GEV) and Gaussian distributions, respectively.

30. a) Regional (3° - 10°), and b) teleseismic network (28° - 90°) for the 1992/03/26 Pahute Mesa, NTS underground nuclear explosion (star). Both the regional and teleseismic

networks are heavily unbalanced, with concentrations of stations in Cascadia and the LA basin (regional), and Europe, Japan and Alaska (teleseismic).

31. Regional Monte Carlo location experiment with the 1992/03/26 Pahute Mesa explosion with a) Gaussian, and b) GEV residuals. *True_iid* (dotted line): independence assumption with independent errors; *False_iid* (dashed-dotted line): independence assumption with correlated errors; *False_cid* (dashed line): correlated assumption with independent errors; *True_cid* (solid line): correlated assumption with correlated errors. The thin grey line represents the cumulative distribution of the coverage parameter for the expected 90th percentile χ^2 distribution with 2 degrees of freedom.

32. Teleseismic Monte Carlo location experiment with the 1992/03/26 Pahute Mesa nuclear explosion with a) Gaussian, and b) GEV residuals. The legend is the same as in Figure 31.

33. Ratios between coverage and mislocation results at each percentile when GEV and Gaussian error distributions were used as Monte Carlo inputs. a) Regional; b) teleseismic case. The differences due to non-Gaussian errors become significant at the highest percentile levels.

34. Regional networks (3°-10°) for the a) 1988/07/07 Pahute Mesa, b) 1988/10/13 Yucca Flat, and c) 1992/11/02 Switzerland GT0 explosions.

35. Cumulative coverage parameters from synthetic Monte Carlo location experiments with a) Gaussian, and b) GEV errors. The symbols get increasingly larger for the 10, 20, 30 and 40-station random sub-networks.

36. Cumulative distributions of mislocation and coverage for GT0-2 nuclear explosions relocated with a) Pn phases between 0°-20°, b) P between 28°-90°, and c) both Pn and P. Green lines denote results with uncalibrated (iasp91) travel-time predictions assuming independence and 1 second reading errors; blue lines show results with calibrated (CUB2, J362D28) travel-times; red lines represent the relocation results with calibrated travel-times when correlated model errors are accounted for and using mb-based measurement error and bias estimates. Black lines represent the cumulative distribution of the coverage parameter for the expected theoretical χ^2 distribution.

37. Mislocations of Yucca Flat explosions using a) uncalibrated Pn, b) calibrated Pn travel-time predictions, and c) calibrated Pn travel-times accounting for correlated model errors and mb-based delay and variance estimates. d-f are the same as a-c but using both regional Pn and teleseismic P phases. Events are color-coded by their mb and locations are plotted relative to the GT0 locations in Easting-Northing coordinates. Error ellipses at the one and two sigma levels encompassing the point clouds are also shown. Black lines represent the cumulative distribution of the coverage parameter for the expected 90th percentile χ^2 distribution with 2 degrees of freedom.

38. Cumulative distributions of mislocation and coverage for GT5 earthquakes relocated with a) Pn phases between 0°-20°, b) P between 28°-90°, and c) both Pn and P. Green lines denote results with uncalibrated (iasp91) travel-time predictions assuming independence and 1 second reading errors; blue lines show results with calibrated (CUB2, J362D28) travel-times; red lines represent the relocation results with calibrated travel-times when correlated model errors are accounted for and using mb-based measurement error and bias estimates. Black lines represent the cumulative distribution of the coverage parameter for the expected χ^2 distribution with 2 degrees of freedom.

39. Mislocation and coverage for 640 GT0-2 nuclear explosions using teleseismic (28°-90°) P phases. The thin grey line represents the cumulative distribution of the coverage parameter for the theoretical χ^2 distribution with 2 degrees of freedom. Green lines denote the baseline results with uncalibrated (iasp91) travel-time predictions when correlated error structure is ignored. Orange lines represent the results with uncalibrated travel-time predictions when correlated error structure is accounted for. Blue lines are results using calibrated (J362D28) travel-time predictions but assuming independent errors. Red lines are cumulatives when correlated errors are accounted for using calibrated travel-time predictions. Coverage and location are improved when correlated structures are accounted for regardless of calibrated or uncalibrated travel times.

Tables

1. Coefficients for 6 degree polynomial in epicentral distance and mb for delay and standard deviation functions.	20
2. Location algorithm configurations for locating 279 Yucca Flats events.	30
3. Location metrics versus location algorithm configurations of Table 1.	30
4. Cluster metrics due to SSSCs and mb-delay corrections and weighting scheme for events with 6 or more regional Pn readings.	35
5. Truth table for location algorithm assumption versus Monte Carlo synthetic errors.	47
6. Coverage versus Monte Carlo inputs using various random sub-networks (10, 20, 40, and 40 stations), and locator assumptions.	52
7. Median mislocation and actual coverage using combinations of travel-time predictions and location algorithm assumptions.	61

1. INTRODUCTION

The objective of this project was to develop methodologies that improve location uncertainties in the presence of correlated, systematic model errors and non-Gaussian measurement errors. Unmodeled lateral heterogeneity in the Earth introduces systematic, correlated travel-time errors that bias both seismic location and location uncertainty estimates. Furthermore, arrival times are frequently picked late and with increasing variance as the signal-to-noise ratio (SNR) decreases, resulting in skewed, non-Gaussian error distributions. Current location algorithms however, either linearized or non-linear, typically assume that observational errors are unbiased and independent, thus ignoring the effects of spatially correlated travel-time predictions and the systematic delay in phase picks with decreasing event size. Our motivation was to characterize correlated model errors and the dependence of measurement errors on event size, and develop an improved location algorithm that accounts for these effects.

We developed a methodology based on copula theory to robustly estimate variogram models for travel-time error. Using this methodology, we produced generic, transportable variogram models for regional Pn and teleseismic P phases. These models are employed to estimate the network covariance matrix describing the spatial correlation structure of travel-time predictions errors. We developed and validated models of measurement errors that map phase picking delay and variance as a function of SNR, as well as a function of epicentral distance and body wave magnitude considered as a surrogate for SNR. The phase pick delays are implemented as travel-time corrections, while the measurement error variances add to the diagonal of the network covariance matrix to obtain the full, non-diagonal data covariance matrix. Our representation of model and measurement errors assumes that the bulk of 3D velocity heterogeneities in the Earth are accounted for by calibrated travel-time predictions. We have developed a linearized iterative location algorithm that utilizes the full data covariance matrix, hence accounts for the correlated model error structure.

By relocating a large number of GT events (GT0-2 nuclear explosions, and GT5 earthquakes produced by Reciprocal Cluster Analysis (Bondár et al., 2007, 2008) we show that ignoring the correlated error structure leads to rapidly deteriorating error ellipse coverage and location bias with increasingly correlated networks. Calibrated travel-times are responsible for the bulk of location improvements, but they do little to improve error ellipse coverage. While the effect of measurement errors on event locations is small (on the order of 1-2 km), they tend to tighten event clusters by reducing the systematic separation between large and small events. We demonstrate that taking into account the correlated model error structure significantly improves error ellipse coverage, and for unbalanced networks, reduces location bias. We also show that the deteriorating effect of ignoring non-Gaussian error distributions, albeit consistent and non-negligible, is of secondary importance compared to the penalty paid for ignoring the correlation structure in travel-time.

2. BACKGROUND

The error budget for a seismic event location is mostly dominated by model errors (travel-time prediction errors due to unmodeled velocity heterogeneities in the Earth) and measurement (picking) errors. The assumption of independent error processes prevails in most modern seismic location algorithms, despite the fact that the problem arising from inadequate representation of systematic bias has been known to seismologists since the advent of modern instrumental seismology. One of the classic examples is the Longshot nuclear explosion (29 October 1965, Amchitka Island, Alaska). Herrin and Taggart (1968) showed that a large number of arrivals traveling along similar ray paths through the oceanic subducting slab, unmodeled by the 1D velocity model, caused large location bias. The systematic travel-time prediction errors produced a 26-km mislocation, far outside the estimated 90% confidence error ellipse (139 km^2). If unaccounted for, correlated systematic errors result in unrealistic error ellipses with degraded coverage (true locations do not lie within the ellipses) and introduce location bias.

Systematic model errors are caused by unmodeled 3D velocity structure as well as stochastic short wavelength variations in the Earth since the source-receiver paths are never exactly the same. Model errors are often modeled by a zero-mean Gaussian distribution. Pavlis (1986) has pointed out that scaling the error ellipse by the RMS residuals (Flinn, 1965) to account for model errors typically fails, as model errors are non-zero mean and not normally distributed. The mean model error represents the systematic travel-time prediction error. Only after this mean is removed by path-specific travel-time corrections, or solved for as part of the inversion (as in multiple event location techniques), can the model errors be approximated by a zero-mean Gaussian distribution. Recognizing the problem of location bias due to inadequate velocity models, most recent location calibration efforts focused on developing better travel-time predictions, either model based or empirical (e.g. Flanagan et al., 2007; Morozov et al., 2005; Murphy et al., 2005; Myers and Schultz, 2000; Reiter et al., 2005; Ritzwoller et al., 2003; Yang et al., 2004).

However, most location algorithms, either linearized or non-linear, assume that the observational errors are all independent. This assumption is violated when the separation between stations is less than the scale length of local velocity heterogeneities. To illustrate the effect of the correlated error structure on seismic event location, we performed a constrained bootstrapping (Yang et al., 2004) experiment on the 7 October 1994 Lop Nor, China nuclear explosion. The explosion is considered GT1 (Fisk, 2002) and recorded by some 400 stations at teleseismic distances. In the experiment we used travel-time predictions obtained from the global 3-D model J362D28 (Antolik et al., 2003) for P phases between 28 and 90 degrees, and located the event with increasing number of stations. At each step we identified the 20 unique subnetworks that provided near-optimal azimuthal coverage for the selected number of stations (starting from 6 to 400) and calculated the median mislocation, area of error ellipse and actual coverage. As **Figure 1a** indicates, the station distribution is far from azimuthally uniform; and is dominated by the networks in California, Japan and Europe. **Figure 1b** shows the trajectory of the mislocation vector with increasing number of stations. As more and

more stations contribute to the solution the location is driven away from the GT1 location. Since the location algorithm does not account for correlated travel-times along similar ray paths, the relative importance of the Californian station clusters steadily increases, resulting in ever increasing location bias.

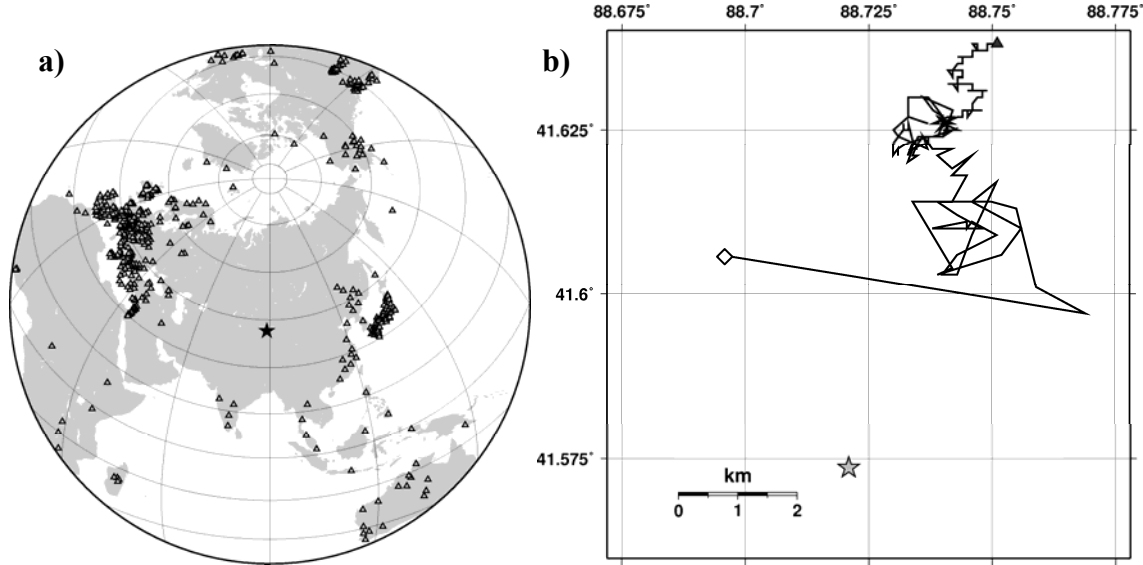


Figure 1. a) Teleseismic stations ($28^{\circ} - 90^{\circ}$) that recorded the 7 October, 1994 underground nuclear explosion at the Lop Nor, China test site. b) Trajectory of median mislocation using subnetworks starting with 6-station networks (open diamond) and gradually increasing to 400 stations (solid triangle). As the number of stations in the subnetworks with optimal azimuthal coverage increases, the location is driven away from the ground truth (star) location.

Figure 2 shows that the size of the error ellipse monotonically decreases with increasing number of stations because the more data (assumed to be independent) are used in the location the smaller the location uncertainties will be. As the information carried by the network geometry is exhausted relatively early, adding more stations merely increases data redundancy. Hence, it is guaranteed that the error ellipse will no longer cover the true location once a sufficiently large number of correlated systematic errors contribute to the solution. Note that we define the coverage parameter (adjusted to GT uncertainty) as

$$\kappa^2 = x^2 / (s_{majax}^2 + GTx^2) + y^2 / (s_{minax}^2 + GTx^2) \quad (1)$$

where x and y are the Cartesian coordinates of the GT event in the coordinate system defined by the semi-major and semi-minor axes of the error ellipse scaled to the 90% confidence level and centered on the seismic location. If the true location is not covered by the 90% error ellipse, the coverage parameter is larger than unity. Under the assumption of independent Gaussian errors, the coverage parameter follows a χ^2 distribution with 2 degrees of freedom. The observation that formal estimates of location uncertainty are often unreliable (i.e. the 90% error ellipses do not cover the true locations 90% of the time) led Bondár et al (2004a) to develop criteria that assess epicenter accuracy independently from the formal error ellipse.

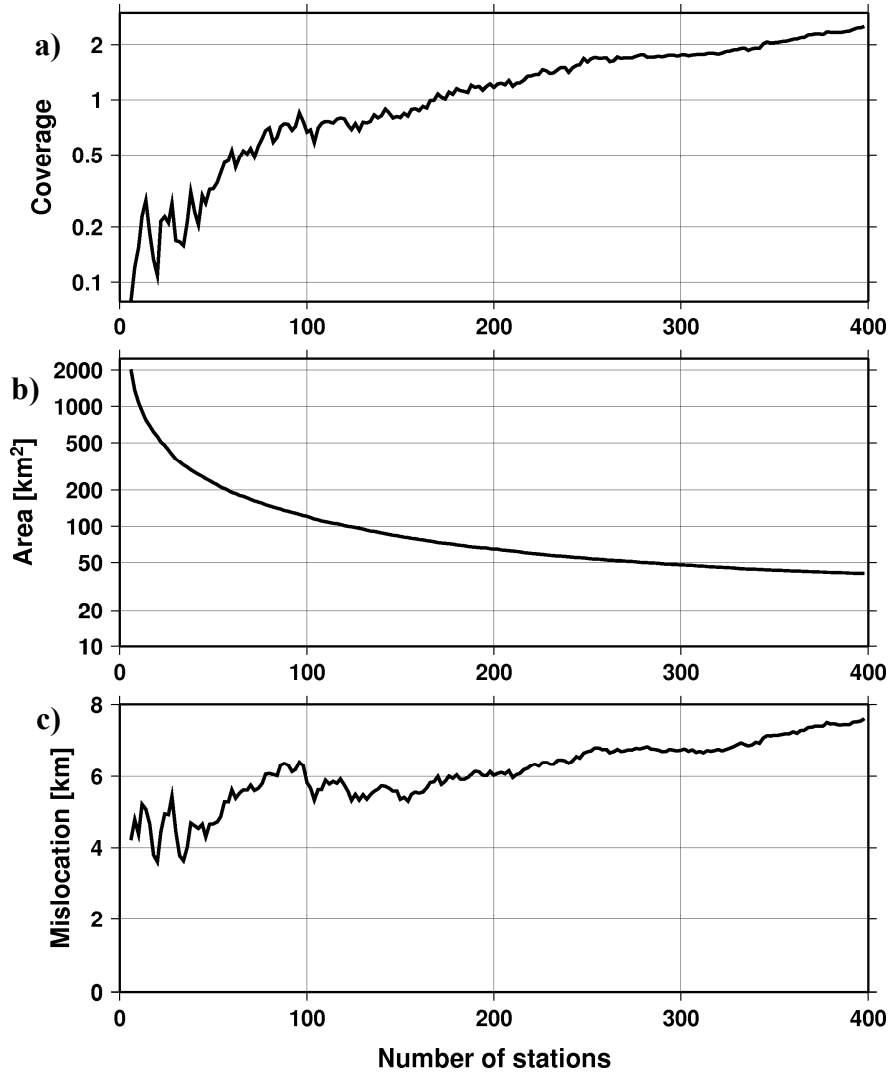


Figure 2. a) Error ellipse coverage, b) error ellipse area and c) mislocation with increasing number of stations for the 7 October, 1994 Lop Nor explosion. Because of the correlated model error structure, the location moves away from the true location, the area of the error ellipse monotonically decreases and the error ellipse no longer covers the true location (coverage > 1) with increasing number of stations.

It has long been noted that errors in P arrival time readings depend on signal-to-noise ratio, SNR (e.g. Freedman, 1966; Lomnitz, 1995). Arrival times tend to be read late and have increasing variance with decreasing SNR (Douglas et al., 2005a). Lateness in readings is not accounted for in standard location algorithms. However, arrival times for a given event read at stations with varying SNRs may, as Douglas et al. (2005b) point out, contribute to bias in the epicentral estimate. Delays in arrival readings may have not been accounted for simply due to lack of reported SNR, since SNR for analogue recordings is difficult to quantify. Indeed, as Douglas et al. (2005a) point out, few actual estimates of reading error characteristics - delay, variance, distributions - have been published. Digital recording with associated automatic processing affords opportunities to characterize reading errors as a function of SNR as well as other signal attributes. For example, the International Data Center (IDC) employs in its location procedures *a priori*

variances for arrival time errors that are a function of SNR estimated by the detector. The IDC SNR is defined as the ratio of the short-term average and the long-term average (sta/lta) of the detecting beam. The IDC processing makes, however, no assumption about *lateness* (or *bias*) in arrival times. Arrival times are assumed to belong to a zero-mean Gaussian distribution.

Owing to recent location calibration efforts, the travel-time model errors that previously dominated the error budget are now often comparable to the measurement (picking) errors. Hence, to achieve further improvements in location accuracy and obtain more reliable location uncertainty estimates, it is necessary to revisit the estimation of these reading errors and develop improved measurement error models for location and uncertainty estimation.

Figure 3 attests to the presence of reading error bias in Pn picks from underground nuclear explosions at the Yucca Flat, Nevada Test Site (NTS). **Figure 3a** shows the iasp91 (Kennett and Engdahl, 1991) GT residuals (i.e. the difference between the observed arrival times and the iasp91 travel time predictions with respect to the GT0 origin times and locations of the explosions) as a function of magnitude, mb. Each circle represents the median residual at a particular station. The two separate populations of median station residuals (one with small, less than a second, and the other with large, 3-4 seconds median residuals) in **Figure 3a** reflect the fact that travel times are fast to the East and slow to the West, and that the global one-dimensional average iasp91 model does not account for the strong velocity heterogeneities observed in North America. The GT0 residuals using predictions from the CUB2 global three-dimensional upper-mantle model (Shapiro and Ritzwoller, 2004) are shown in **Figure 3b**. Once the bulk of systematic path effects are accounted for by the CUB2 model, a clear trend emerges, indicating increasingly late picks with decreasing magnitude. Note that using calibrated travel times reveals the measurement errors.

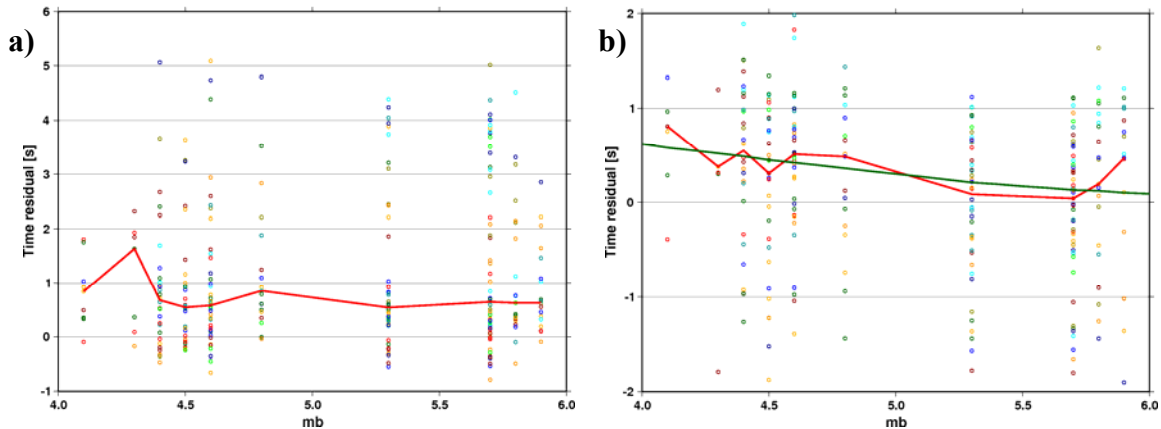


Figure 3. a) Pn GT0 residuals (observed - iasp91 predictions with respect to the GT0 locations and origin times) for explosions at Yucca Flat, Nevada Test Site plotted as a function of magnitude, mb. Circles represent the median residuals at individual stations. The red line connects the median station residuals. The non-zero median indicates that the iasp91 velocity model is too fast for the Western US. b) Pn GT0 residuals from CUB2 model predictions. The CUB2 model reduces the path effects not modeled by iasp91 and reveals a trend (green) due to systematically late picks with decreasing event size. The median bias decreases by 0.3 seconds from mb 4.5 to 5.5.

Variations in SNR may bias not only epicenters but also empirical estimates of path effects or corrections to global travel time models such as iasp91. These path effects are also being referred to as source-specific station corrections or SSSCs (e.g. Morozov et al., 2005; Murphy et al., 2005; Ritzwoller et al., 2003; Yang et al., 2001, 2004).

3. TECHNICAL APPROACH

3.1. Location algorithm with correlated errors

The promise of developing improved travel-time predictions, typically implemented as travel-time correction surfaces with respect to a background 1D velocity model, is that they not only improve locations but also reduce model errors due to unmodeled velocity heterogeneities. Indeed, perfect travel-time predictions would carry no model errors, and predictions along very similar ray paths would remain uncorrelated. Unfortunately, 3-D velocity models are never perfect, and one never has enough data points to obtain perfect empirical travel-time corrections through kriging (e.g. Morozov et al., 2005; Myers and Schultz, 2000). Hence, it is advisable to account for the correlated nature of travel-time predictions for waves traveling along similar ray paths. When correlated model errors are present, the data covariance matrix is no longer diagonal, and the redundancy in the observations reduces the effective number of degrees of freedom. Thus, ignoring the correlated error structure inevitably results in underestimated location uncertainty estimates. For events located by an unbalanced seismic network (either sparse or dense) this may also lead to biased location estimates.

3.1.1. Model errors: Estimation of the network covariance matrix

Chang et al (1983) showed that incorporating correlated error structure into a linearized location algorithm is relatively straightforward. They reported that introducing a rather simple *a priori* correlation matrix into the location algorithm reduced the location error of the Longshot explosion to 3.2 km with the error ellipse covering the true location. The real difficulty however, is to obtain a robust estimate of the full covariance matrix itself. Recently Rodi and Myers (2007) developed a model-based approach to compute travel-time residual covariance matrices. Their method is basically formulated as a single-event tomography problem, where the covariance matrix is obtained by integrating travel-time sensitivity kernels in a 3-D velocity model with an *a priori* correlation function. In their representation the full covariance matrix depends on both station and event locations, thus allowing for non-stationary correlations.

We follow a different, empirical approach to estimate the full covariance matrix. We assume that station separation is a good approximation to measure ray path similarity. This simplifying assumption allows us to avoid expensive ray tracing through 3-D models and to estimate covariances between station pairs from a variogram model. Because in this representation the covariances do not depend on event locations, the covariance matrix (and its inverse) needs to be calculated only once.

In order to obtain robust variograms, we use entire clusters of ground truth events. The advantage of using event clusters is that similar ray paths are sampled multiple times which mostly eliminates the contaminating effect of picking (measurement) errors. Furthermore, using GT5 or better events (besides the obvious advantage of not being affected by gross location errors) allows us to calculate ‘ground truth’ residuals, that is, residuals calculated with respect to the GT location, using a specific Earth model.

The isotropic, stationary variogram in geostatistical analysis is defined by the equation

$$\gamma(h) = \left\langle (\delta t_i - \delta t_j)^2 \right\rangle = \sigma_{sill}^2 - Corr(h)\sigma_{sill}^2 = \sigma_{sill}^2 - Cov(h) \quad (2)$$

where σ_{sill}^2 denotes the background variance, δt_i and δt_j are the ground truth residuals at station pairs of common events, and $h = \Delta(sta_i, sta_j)$ is the station separation, and $Corr$ and Cov are the correlation and covariance matrices, respectively. Instead of forcing a parametric variogram model (such as Gaussian, spherical or exponential) typically used in geostatistics, we define the robust variogram model as the median regression curve of the residual difference squares for station pairs of common events with respect to station separation. Copula theory (e.g. Nelsen, 1999; Salvadori et al, 2006) offers an elegant and fully data-driven approach to determine this median regression curve. We describe the copula formalism for the robust estimation of the variogram in the **Appendix**. Having derived a variogram model, the estimates for the elements of the network covariance matrix, which characterizes the model errors, are obtained as

$$C_N(i, j) = \sigma_{sill}^2 - \gamma(\Delta(sta_i, sta_j)) \quad (3)$$

Note that the isotropic variogram model does not account for azimuthal variations in the correlation structure. Hence, the proper choice of the underlying velocity model to calculate travel-times becomes important, especially for regional phases. To account for the bulk of 3D velocity heterogeneities in the Earth we use the global upper-mantle CUB2 (Shapiro and Ritzwoller, 2004) and the global whole-mantle J362D28 (Antolik et al., 2003) 3-D models to obtain travel-time predictions for regional and teleseismic phases, respectively. Both models have been shown to improve seismic locations and achieve significant variance reduction in travel-time predictions (e.g. Bondár et al, 2004b; Ritzwoller et al., 2003; Yang et al., 2004).

Figure 4 shows the GT event clusters we use throughout this study. The data set consists of GT0-2 nuclear explosions from the SAIC Nuclear Explosion Database (www.rdss.info, Bahavar et al., 2007), and GT5 earthquakes produced by the hybrid Hypocentroidal Decomposition and Reciprocal Cluster Analysis (HDC-RCA, Bondár et al., 2007, 2008). We use arrival times and phase names from the groomed ISC bulletin (EHB, Engdahl et al., 1998), which for the earthquake clusters were subjected to further quality control during the HDC-RCA processing. We use only first arriving Pn or P phases. Note that most of the outliers were removed from the data set during the EHB and HDC-RCA analyses, thus, the Pn/P arrival data represents a clean, quality controlled data set.

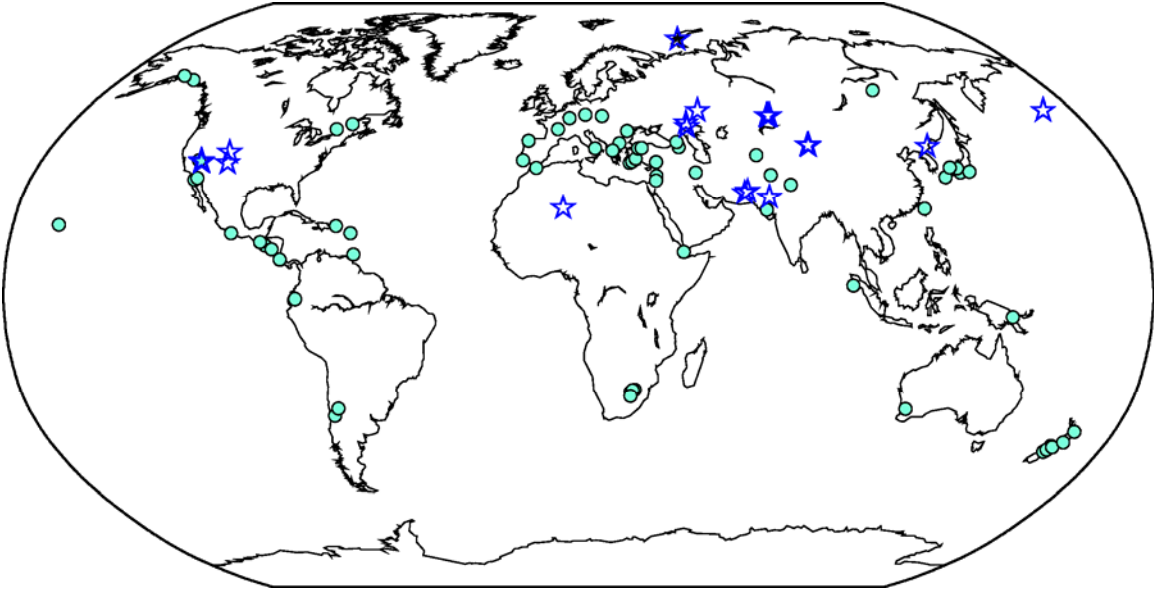


Figure 4. Globally distributed GT5 earthquake clusters (circles) and GT0-2 clusters of nuclear explosions (stars).

Figure 5 shows the Pn and P variogram models for each event cluster obtained from GT residuals using the CUB2 and J362D28 travel-time predictions. The fact that the variograms look quite similar indicates that the CUB2 and J362D28 global 3D models indeed account for major 3D heterogeneities. Consistent with previous results the calibrated travel-times reduced the overall variance (the sill) by nearly 50%. Nevertheless, unmodeled velocity structures remain that generate correlated travel-time residuals. For instance, the two Pn variograms with distinctly larger sill than the rest were derived from the Racha and Spitak clusters, both situated in the Caucasus, indicate that smaller-scale velocity heterogeneities in the region still remain unexplained by the CUB2 model.

Because the calibrated travel times account for the bulk of 3D Earth structure, it allows us to derive an isotropic variogram model not just from the individual clusters, but from the entire data set of globally distributed GT5 or better events. These models, shown as thick red lines on **Figure 5**, define generic Pn and P variograms that can be used to construct network covariance matrices for regional and teleseismic networks. The generic variogram models may be too optimistic for some regions and may be too pessimistic for others, but because the variability of the variograms derived from the individual clusters is relatively small, we expect that the generic variogram models will perform reasonably well anywhere in the globe. This notion is especially important for regions with insufficient data to derive a variogram model.

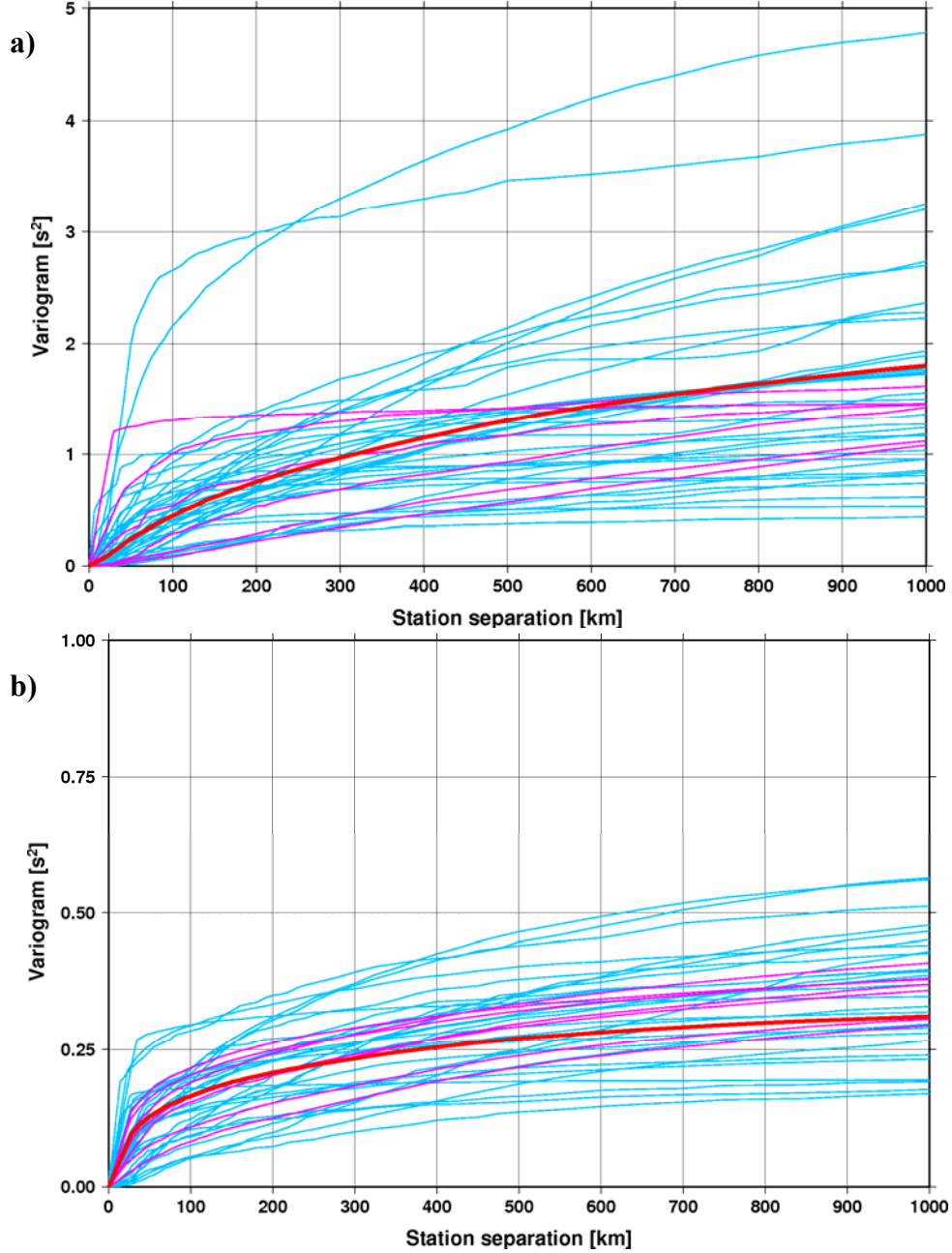


Figure 5. Variogram models for a) regional Pn and b) teleseismic P derived from earthquake (blue) and explosion (magenta) clusters. The thick red lines show the generic variogram models derived from all clusters.

As we mentioned above, the network covariance matrix derived from the generic variogram models represents the error budget due to model errors. For the background model error variance we have found that sill values of $\sigma_{sill} = 1$ second for teleseismic P and $\sigma_{sill} = 1.53$ seconds for regional Pn provide adequate coverage. We consider the picking errors as independent, Gaussian variables. Hence, to obtain the full data covariance matrix, the variances of the measurement errors simply add to the diagonal of the network covariance matrix.

3.1.2. Measurement errors: Bias and variance

In this study SNR dependence of both delay and variance of reading errors of first arriving P waves are analyzed and functional relations for delay and variance are estimated. The analysis includes data from the data set described above, as well as data from PIDC/IDC Reviewed Event Bulletins (REB). We resort to amplitude/period ratios and network mb magnitudes reported in the ISC bulletins (www.isc.ac.uk) as surrogates for directly measured SNR (Douglas et al., 2005b). For the IDC REB data, the detector sta/lta estimate is calculated consistently for all arrivals and it is a direct measure of SNR.

Investigating reading errors is facilitated by observations from ground truth (GT0) events, i.e., events for which locations as well as origin times are known precisely. Furthermore, observations from suites of events at the same location with varying magnitude make it possible to map dependence on SNR. This analysis focuses on P arrivals from NTS explosions, a data set which fulfills the requirement of ground truth and range of event size. For other datasets we turn to estimates based on double-differences between arrival times of station pairs which mitigate the uncertain earthquake or explosion origin times.

3.1.2.1. NTS data

The data include first arrival times of underground nuclear explosions between 1965 and 1987, 65 events at Pahute Mesa and 274 events at Yucca Flat. The explosions cover a wide range of network magnitudes ($4.0 < mb < 6.7$) and were recorded at both regional and teleseismic distances. The locations and origin times reported by Springer et al. (2002) provide ground truth for the explosions. The arrival times associated with the explosions were limited to first P arrivals at stations within 90 degrees from NTS. Requiring at least 50 recorded events from the two sites combined at each station, there were arrival data at 200 stations, located primarily in North America at regional and teleseismic distances, and in Europe at teleseismic distances. Röhme et al (1999) pointed out that the quality of arrival picks reported to ISC vary vastly among stations due to different analyst and processing procedures as well as recording equipment. As the data cover about 30 years there are also time dependencies in the measurements reflecting three decades of changes in personnel, procedures, and instrumentation.

The GT0 arrival time residual is defined as the difference between the measured arrival times and the iasp91 travel-time prediction with respect to the GT0 ground truth origin time and hypocenter. The arrival time residual is thus the sum of the true path effect (or true SSSC) and the measurement error where the true path effect is the difference between the true travel time and that calculated from the iasp91 travel-time model. If we assume the same path effect for a given station and NTS test area (Pahute Mesa or Yucca Flat) the residuals describe the variation in measurement error. We write $r_{ij} = p_i + e_{ij}$ where r_{ij} is the residual at station i from explosion j , p_i is the path effect for station i and e_{ij} is the measurement error at station i from explosion j .

There are several sources of variability. Douglas et al. (2005a) estimate that the range of variation of the difference between true and predicted travel times for NTS explosions

is about 0.2 seconds due to variation in height above sea level. A rounding error in reporting precision will introduce scatter in the residuals r_{ij} . Many arrival times were reported to the nearest 0.1, 0.5 or 1.0 second. For example, the standard deviation of 1 second uniform rounding error is 0.3 second. A review of the distribution of the decimals of reported arrival times showed that none of the stations consistently reported with a single precision. With reduced SNR the reading accuracy can be expected to go down. The expected distribution of the tenth of the second of the true arrival times for arrivals at a given station from closely spaced explosions may not be uniform. In this paper we made no attempt to consider variations in reported precision.

3.1.2.2. *Measurement error model based on amplitude/period ratio*

We analyze below the residuals, r_{ij} , as a function of the logarithm of station amplitude/period ratios A/T , and as a function of network mb . As the background noise at a given station varies from explosion to explosion, we expect A/T on average to be directly proportional to the SNR. The standard deviation of ambient noise amplitudes is typically 0.2 (log10 scale A in nm), but can be larger for some stations with strong seasonal variations such as YKA (Yellowknife array, Canada) (Douglas et al, 2005c). As most explosions were typically detonated in a fairly narrow time window of the day, diurnal variation in noise is expected to have a small effect. Change in dominant frequency of recorded signals with event strength probably also contributes to deviations from a direct proportionality between A/T and SNR. Sufficient A/T data were available for 61 stations to test this hypothesis.

Figure 6a shows the arrival time residuals at station COL (College Outpost, Alaska), at an epicentral distance of 32° , as a function of $\log A/T$ for explosions from Pahute Mesa and Yucca Flat. For Yucca Flat events (solid circles), the residuals become on average increasingly delayed with decreasing A/T . These residuals are fairly constant for A/T above about 40 nm/s, but they increase by about 1 second when A/T decreases from 40 to 10 nm/s. The Pahute Mesa residuals (open circles) are almost constant throughout their range (which is well above 40 nm) and there is no striking delay effect with decreasing A/T values.

The delay of the residuals tends to plateau above some A/T value. As the A/T becomes sufficiently large, residuals converge to a constant value and the delay in picked arrival times becomes negligible. This constant value is an estimate of the path effect. A positive value of path effect means that the travel time along the source to station path is slower than predicted by the iasp91 travel time model.

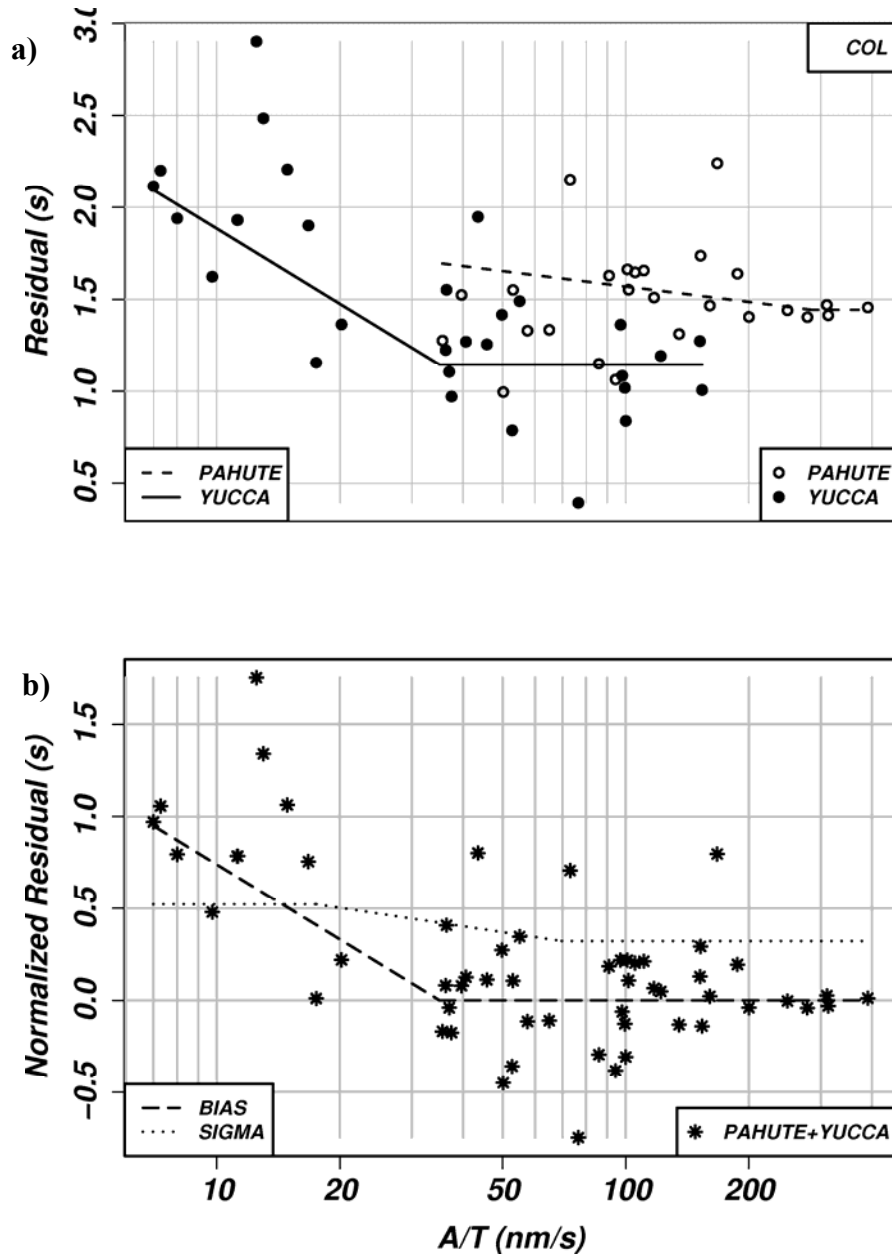


Figure 6. a) Arrival time residuals as a function of amplitude/period ratios (A/T) for the station COL (College Outpost, Alaska) for NTS underground nuclear explosions. Different symbols are used for explosions at the two areas Pahute Mesa (open circles) and Yucca Flat (filled circles) of NTS. There is a systematic difference in arrival time residuals between the two areas as indicated by the estimated relations (solid and dashed lines) between arrival time residual as a function of A/T. b) The residuals for Pahute Mesa and Yucca Flat in Figure 2) have been combined after subtracting estimated path effects. The estimated bias and standard deviation in measurement error as a function of A/T are indicated as dashed and dotted lines.

Given the model suggested by **Figure 6a**, we fit piecewise linear trend-lines to GT residuals. Trend-lines of the delay, d_i , at station i , for a given event area are defined as

$$d = p_i \text{ for } sn \geq sn_b \quad (4)$$

$$d = p_i + d_0(sn - sn_b)/(sn_b - sn_0) \text{ for } sn_0 \leq sn \leq sn_b \quad (5)$$

where p_i is the path correction, $sn = \log_{10}(A/T)$, $sn_0 = \min(\log_{10}(A/T))$, and sn_b is the minimum value of $\log_{10}(A/T)$ above which the delay, d is constant. The difference, $d_i - p_i$, represents the delay in measurement error. This means that an estimated trend-line consists of two linear segments, a horizontal line for the largest A/T values, which at some A/T value joins a line with negative slope. The trend-lines were fitted with the L1 norm and were usually close to estimates based on a locally weighted regression algorithm (Cleveland, 1979). Estimates based on the analytical form of the trend-lines are preferred as they are always monotonic.

The estimated path effects of Pahute Mesa and Yucca Flat for COL in **Figure 6a** are slightly offset. The Yucca Flat path effect is about 0.2-0.3 second smaller than that for Pahute Mesa, suggesting a slightly faster travel path from Yucca Flat. In **Figure 6b** the Pahute Mesa and Yucca Flat residuals have been combined after normalization by subtracting the estimated path effects. The estimated trend line for the combined data is shown as a dashed line and represents the delay due to measurement error. The standard deviation of the measurement error along the trend-line was obtained as the *smad* (standard median absolute deviation) estimated from a moving data window and is marked as a dotted line in **Figure 6b**.

In **Figure 7**, residuals as a function of A/T at COL and EDM (Edmonton, Canada) are compared for Yucca Flat explosions. Also marked, as vertical lines, are the A/T thresholds (50%) estimated by Lilwall and Neary (1986) based on amplitude/period data from seismic events with a broad global distribution. The Lilwall and Neary (1986) thresholds are consistent with the observed A/T data for the two stations as only few observations are below the thresholds. We can normalize the A/T values and the residuals for a given station by subtracting the Lilwall and Neary (1986) threshold and the estimated path effect. Residuals for Pahute Mesa and Yucca Flat events are combined by subtracting separate estimates of the path effects. Then, by subtracting the Lilwall and Neary (1986) station threshold from the $\log(A/T)$ values we obtain a normalized measure of SNR (or $\log_{10}(\text{SNR})$) relative to the Lilwall and Neary (1986) station threshold. Residuals and A/T values normalized in this manner can then be used for comparison of picking errors between stations.

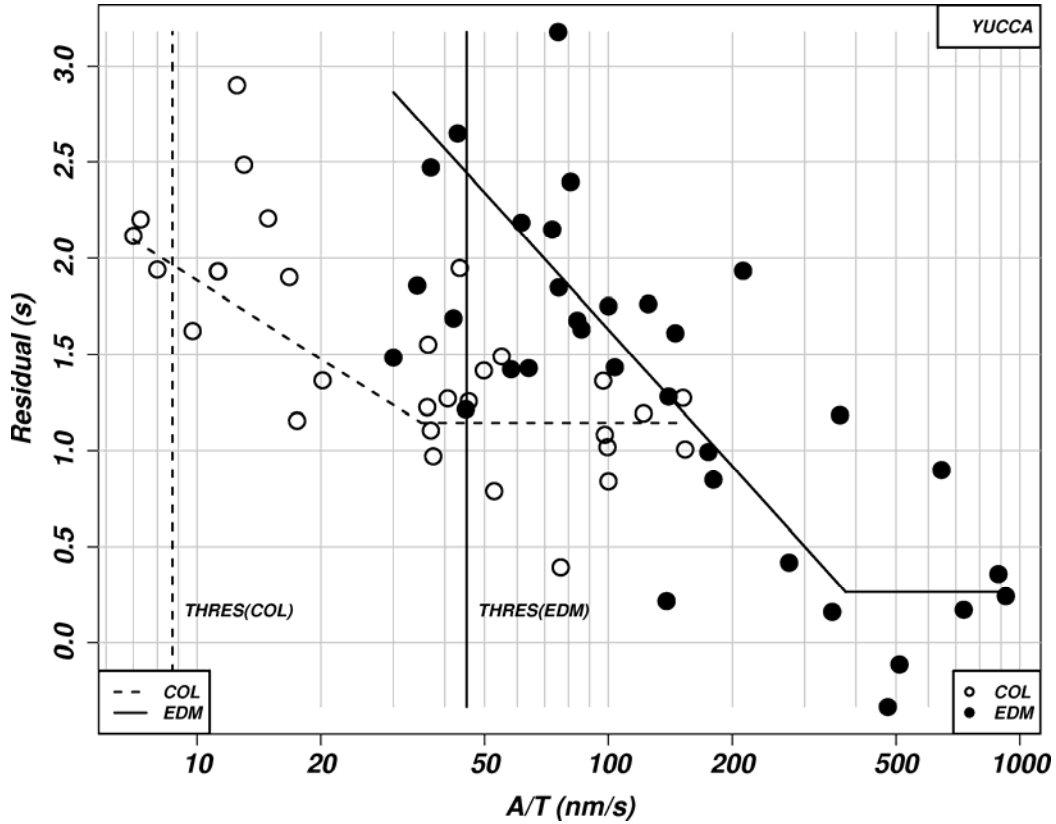


Figure 7. Arrival time residuals as a function of amplitude/period (A/T) ratios at the two stations COL (College Outpost, Alaska) and EDM (Edmonton, Canada) for explosions at Yucca Flat. Dashed and solid lines mark the thresholds from Lilwall and Neary (1986) for COL and EDM respectively.

3.1.2.3. Measurement error model based on mb and epicentral distance

Stations that report manual readings usually do not report A/T for all detected signals, and many operators do not report amplitudes. This limits the availability of A/T measurements to large events, especially for historical data. Current analysis and reporting procedures, such as those employed by the IDC, do, owing to digital recordings and processing, measure and report A/T data for each detected phase. For historical data network mb is an alternative to A/T as a measure of SNR. Network mb is expected to have less correspondence to SNR than A/T; for example, empirically the scatter in station mb has been estimated to be 0.35 (Veith and Clawson, 1972) While admittedly a relatively poor surrogate for SNR, we may still expect the logarithm of the SNR proportional to mb. Network mb not only provides a much larger data set, but also a broader range of magnitudes. There were arrival time data available for 50 or more events with network mb from about 200 stations compared to only 61 stations with A/T data.

Figure 8a shows the time residuals as a function of network mb for the station COL. Relations for delays and standard deviation of the residuals as a function of network mb were estimated in a similar way as for A/T above. Path effects were estimated from the horizontal trend-lines at large magnitudes. After subtracting estimated path effects,

residuals were combined and used to estimate delay and scatter as a function of mb (Figure 8b). The trends of the data as a function of mb for the station COL are similar to those of A/T in Figure 2. The data for mb extends to smaller events than A/T data, thus spanning a wider range in surrogate SNR. However, the scatter of the residuals around the trends appears, as expected, somewhat larger.

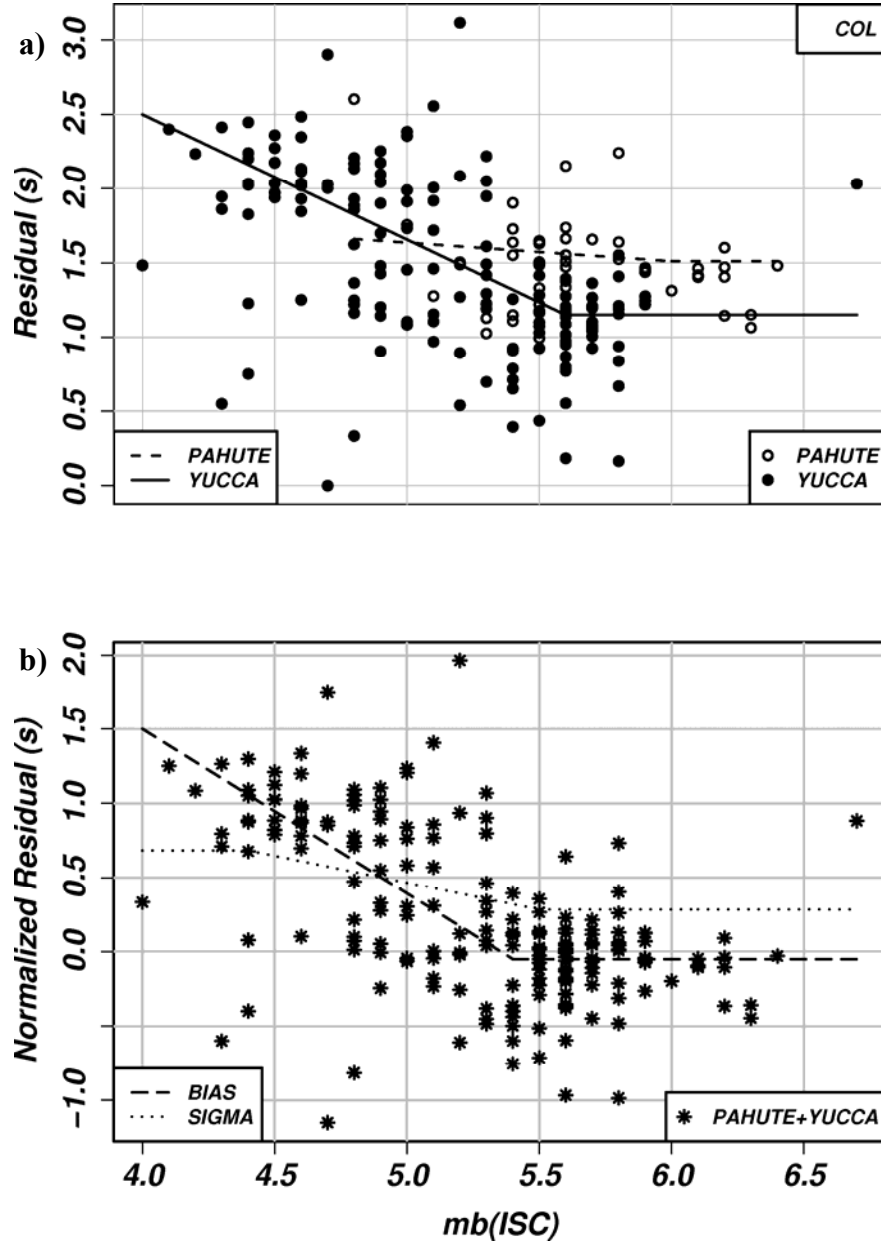


Figure 8. a) Arrival time residuals as a function of network mb, mb(ISC), for the station COL (College Outpost, Alaska) for NTS underground nuclear explosions. Different symbols are used for explosions at Pahute Mesa (open circles) and Yucca Flat (filled circles). There is a systematic difference in arrival time residuals between the two areas as indicated by the estimated relations (solid and dashed lines) between arrival time residual as a function of mb(ISC). b) Pahute Mesa and Yucca Flat data combined after normalization. The estimated bias and standard deviation in measurement error as a function of network mb are indicated as dashed and dotted lines.

Estimated delays and standard deviations (sigma) as a function of network mb and as a function of $\log A/T$, normalized with the Lilwall and Neary (1986) thresholds ($SNR = 1$ at the threshold), are compared for 60 stations in **Figure 9**. Delay and standard deviations of residuals were corrected for path effects. For A/T (**Figure 9b**) the delay for most stations becomes non-negligible for $SNR < 10$ times the threshold; the delay for most stations is about 0.3 second near the threshold, but there are stations with delays larger than 1 second at the threshold. The estimated standard deviations as a function of SNR show quite large variation. Generally the delay functions for network mb (**Figure 9a**) and for A/T for a given station are similar, whereas the sigma functions for network mb generally are larger (typically by a factor of 2). The range of network mb is typically one order of magnitude larger than that based on $\log_{10}(A/T)$ for most stations. Regional stations tend to have the largest delays and largest standard deviations. Note that differences in procedures among stations to read arrival times may also contribute to the variation in estimated delay and sigma functions. The station specific delay and standard deviation functions corrected for path effects characterize reading errors of arrival times that can be used in the location of NTS events. In a location algorithm arrival times can be corrected for delays calculated from the delay functions and the arrival time data can be weighted with the standard deviations calculated from the sigma functions.

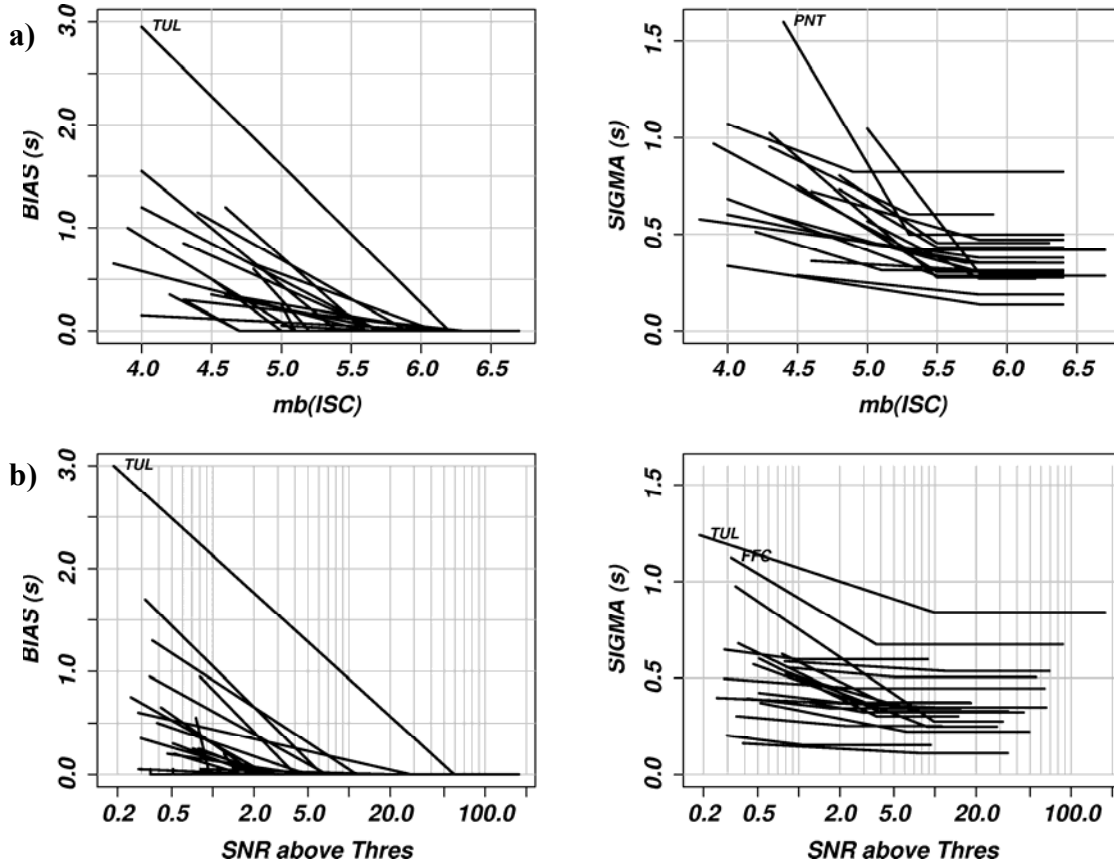


Figure 9. Estimates of delay (BIAS) and standard deviation (SIGMA) in arrival times of NTS explosions a) as a function of network mb(ISC) for 60 stations, b) as a function of SNR relative to the Lilwall and Neary (1986) thresholds ($SNR = 1$).

The estimated path corrections, delay and standard deviation functions can also be used to compare the distribution of the measurement errors with a Gaussian distribution. Assuming independent measurements the normalized residuals, for all stations, i , and all events, j , then $(r_{ij} - d_i(s_{n_j})) / \sigma_i(s_{n_j})$ would, if Gaussian, all belong to the standard normal distribution with zero mean and unit standard deviation.

The histograms in **Figure 10** compare distributions of residuals for Yucca Flat explosions as corrections are progressively applied. Gaussian density functions with mean and standard deviation of the empirical data are overlaid the histograms. **Figure 10a** shows the histogram for the iasp91 residuals, r_{ij} . It is a skewed distribution with an offset of about 0.5 second. Most of this offset is removed by applying the estimated path corrections $(r_{ij} - p_i)$ as shown in **Figure 10b**. **Figure 10c** indicates that the offset is completely removed when the bias of the measurement errors is corrected for $(r_{ij} - p_i - d_i(s_{n_j}))$. The distribution with measurement error bias removed also becomes slightly more symmetrical, but the distribution still deviates significantly from a Gaussian with its comparatively heavy tails. In a final step the residuals have been normalized with the estimated standard deviation of the measurement errors, $\sigma_i(s_{n_j})$, and are shown in **Figure 10d**. The horizontal scale for these normalized residuals is in sample quantiles. The distribution of these normalized residuals is still not Gaussian with its histogram clearly differing from the density function of a standard Gaussian distribution (zero mean and unit standard deviation), which is drawn as a dashed line. The discrepancy could be due to over-estimated standard deviations; the distribution becomes close to Gaussian if the estimated standard deviations are reduced by 25%.

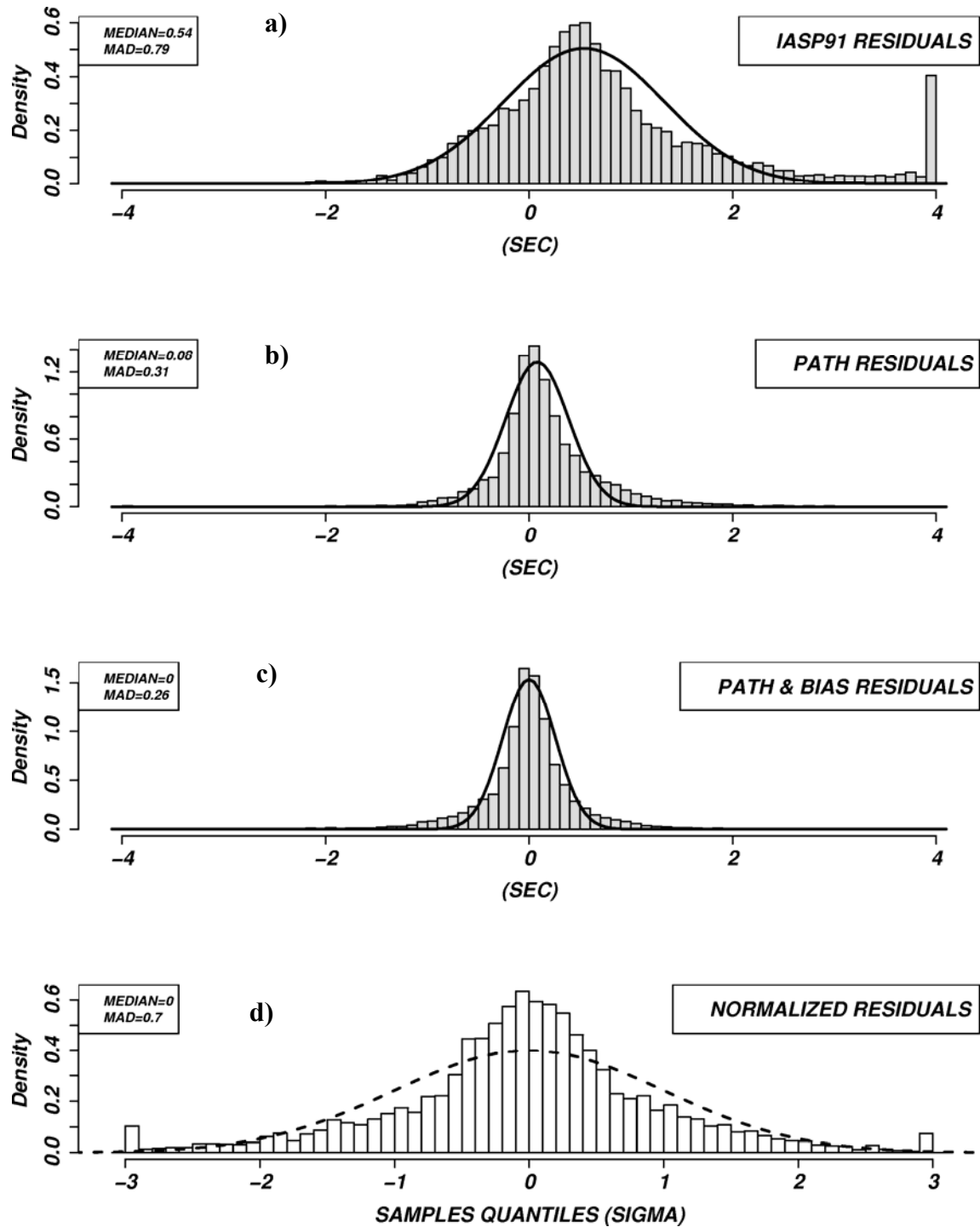


Figure 10. Histograms for arrival time residuals of Yucca Flat events compared with Gaussian distributions. a) iasp91 residuals. b) residuals with estimated path corrections. c) residuals with path corrections and mb-based delay corrections. d) normalized residuals: path and bias corrected residuals are divided by the mb-based standard deviation estimates. Note that the horizontal scale for the normalized residuals is in sample quantiles and the density function of the standard Gaussian (zero mean and unit standard deviation) is drawn as a dashed line.

As delay and sigma functions could not be estimated for about 1,000 stations which recorded fewer than 50 explosions we derived generic delay and sigma relations as functions of network mb and epicentral distance. That is to say correction curves would not be tailored to each individual station. **Figure 11** shows the contours of delay and standard deviation as a function of mb and epicentral distance derived from all the available residual data. We fitted a 6-degree polynomial surface to the contours in order to provide closed expressions for delay and sigma as a function of mb and distance (see Venables and Ripley, 2002). The coefficients of the polynomial are listed in **Table 1**. **Figure 11** suggests that the effect of phase pick delays on event locations with decreasing magnitude is small or negligible for teleseismic networks, but cannot be ignored for local and regional stations.

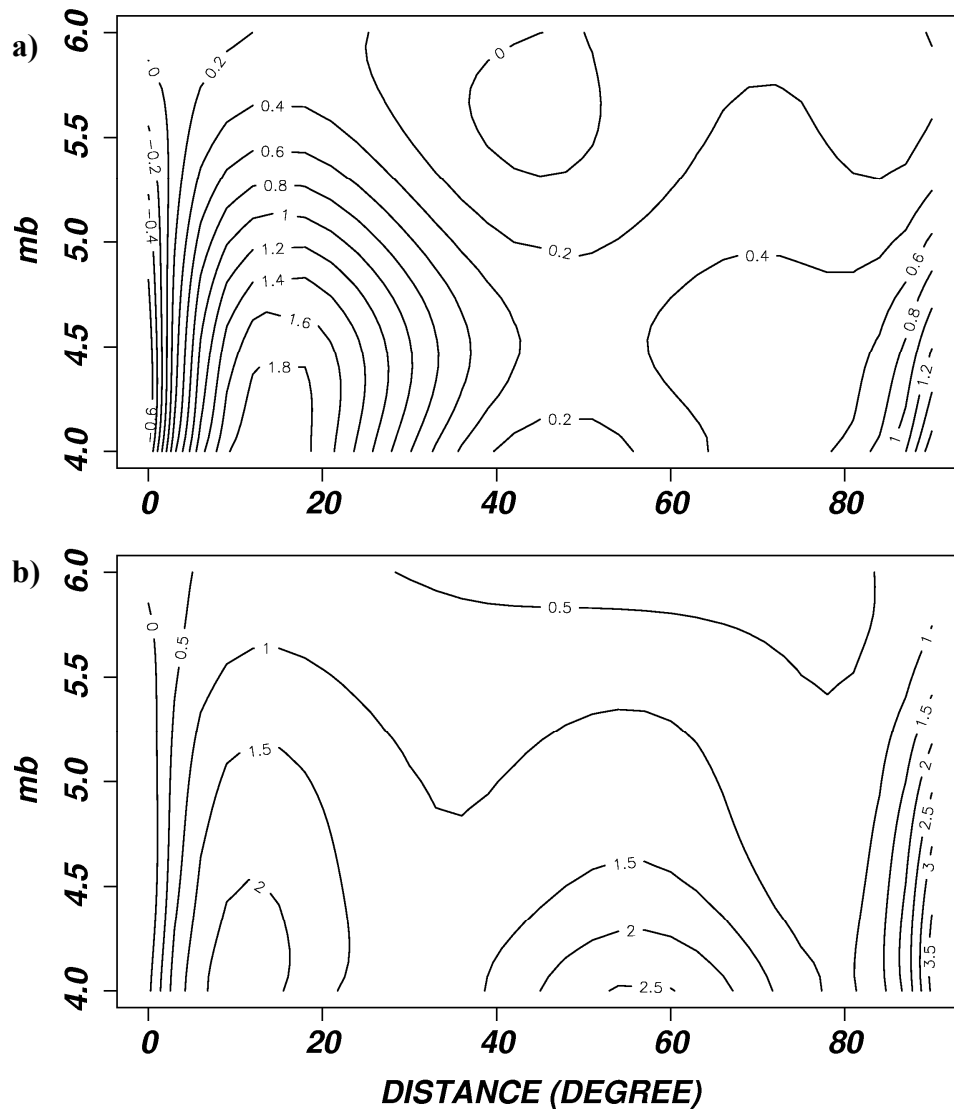


Figure 11. The contours show the a) delay and b) standard deviation in arrival time measurements as a function of network mb and epicentral distance for NTS explosions.

Table 1. Coefficients for 6 degree polynomial in epicentral distance and mb for delay and standard deviation functions. The coefficients are given for normalized distance, $\Delta' = 2\Delta / 93.25 - 1$, and $mb' = 2(mb-3.8) / 2.2 - 1$. The polynomial is defined as $p(\Delta, m_b) = \sum_{N,M} c_{N,M} \Delta'^N m_b'^M$

N	M	Bias Coefficient	Sigma Coefficient
0	0	0.1561	1.1215
1	0	-0.0793	0.9312
2	0	3.1146	-0.8035
3	0	-2.9457	-7.8131
4	0	-4.6197	1.9336
5	0	3.7841	8.9413
6	0	1.6314	-0.4692
0	1	-0.5723	-0.4697
0	2	0.7671	-1.0111
0	3	-1.3986	1.0407
0	4	1.4666	6.8377
0	5	1.3989	-2.2638
0	6	-3.2485	-7.7796
1	1	0.0213	-0.0119
1	2	0.0706	0.6839
1	3	-0.3566	-0.9987
1	4	0.0833	-0.8795
1	5	0.5907	0.7972
2	1	0.8355	-0.7039
2	2	-0.7698	-1.0255
2	3	-0.3802	0.9666
2	4	1.0768	1.6698
3	1	-0.2634	0.4409
3	2	-0.0299	-0.0485
3	3	0.2551	0.2156
4	1	-0.2990	0.2332
4	2	-0.0141	0.0292
5	1	0.1185	-0.3586

3.1.2.4. Measurement error model based on SNR

IDC processing makes consistent *sta/ta* measurements for all detections. As the IDC operations only recently began, there is little GT data with known GT0 origin times to investigate measurement error as a function of SNR. Therefore, we use double-differences of arrival-time residuals from closely spaced and well-located events (GT5 epicenter but no GT5 origin time) to estimate delay and standard deviation as a function of SNR.

We use arrival times with corresponding SNR estimates from a network of stations available from a cluster of GT events. We write the arrival time at station i from event x as

$$t_{ix} = OT_x + t_{ix}^{pred} + t_i^{path} + t_{ix}^{delay} + \varepsilon_{ix} \quad (6)$$

where OT_x is the origin time of event x , t_{ix}^{pred} is the predicted (iasp91) travel-time, t_i^{path} , is the path effect (station term), t_{ix}^{delay} is the deterministic pick error (delay) and ε_{ix} is a random picking error. We assume only t_{ix}^{delay} and ε_{ix} are functions of SNR_{ix} . By forming selected double differences, the origin times and systematic station terms may be canceled:

$$\begin{aligned} (t_{ix} - t_{iy}) - (t_{jx} - t_{jy}) - (t_{ix}^{pred} - t_{iy}^{pred}) + (t_{jx}^{pred} - t_{jy}^{pred}) = \\ (t_{ix}^{delay} - t_{iy}^{delay}) - (t_{jx}^{delay} - t_{jy}^{delay}) + (\varepsilon_{ix} - \varepsilon_{iy}) - (\varepsilon_{jx} - \varepsilon_{jy}) \end{aligned} \quad (7)$$

The right hand side of equation (7) can be calculated from the observed arrival times and the predicted travel-times. The mean delay and the total variance (assuming that the travel-time prediction errors are more or less the same) are written as

$$\begin{aligned} \mu^{delay}(SNR) &= (t_{ix}^{delay} - t_{iy}^{delay}) - (t_{jx}^{delay} - t_{jy}^{delay}) \\ Var^{total} &= 4Var_{hypo} + Var(SNR_{ix}) + Var(SNR_{iy}) + Var(SNR_{jx}) + Var(SNR_{jy}) \end{aligned} \quad (8)$$

For sufficiently large SNR values, we assume the pick delay and variance approach zero. Thus, if all four readings have large SNR, the total variance yields an estimate of the model error variance. This provides an estimate of the variance of reading errors when all four readings have similar SNR: $Var(SNR) \rightarrow (Var^{total} - 4Var_{hypo}^{I \arg eSNR})/4$. Now let us presume that three of the readings have large SNR, and one has a small SNR. Then the mean delay simply becomes $\mu^{delay}(SNR) \rightarrow t^{delay(SNR_{small})}$. Hence, the double-difference approach provides a methodology to derive models of bias and variance in picking errors.

To establish SNR-dependent models of delay and variance, we used first-arriving P from 12 GT5 earthquake clusters (Bondár et al., 2007, 2008) with $SNR \geq 3$ reported in the PIDC/IDC REB. We calculated all possible double-differences (DD) in each cluster and selected two subsets for each cluster. One subset consisted of the DD residuals for which the SNR of all four arrivals were similar within a factor of two. **Figure 12a** shows the DD residuals for this subset as a function of median SNR, as well as the SMAD of residuals along moving SNR windows. The other subset included DD residuals with the three largest $SNR \geq 40$. The DD residuals are plotted as a function of minimum SNR in **Figure 12b**. The running median (dashed curve) is the estimate of delay or bias in reading error as a function of SNR.

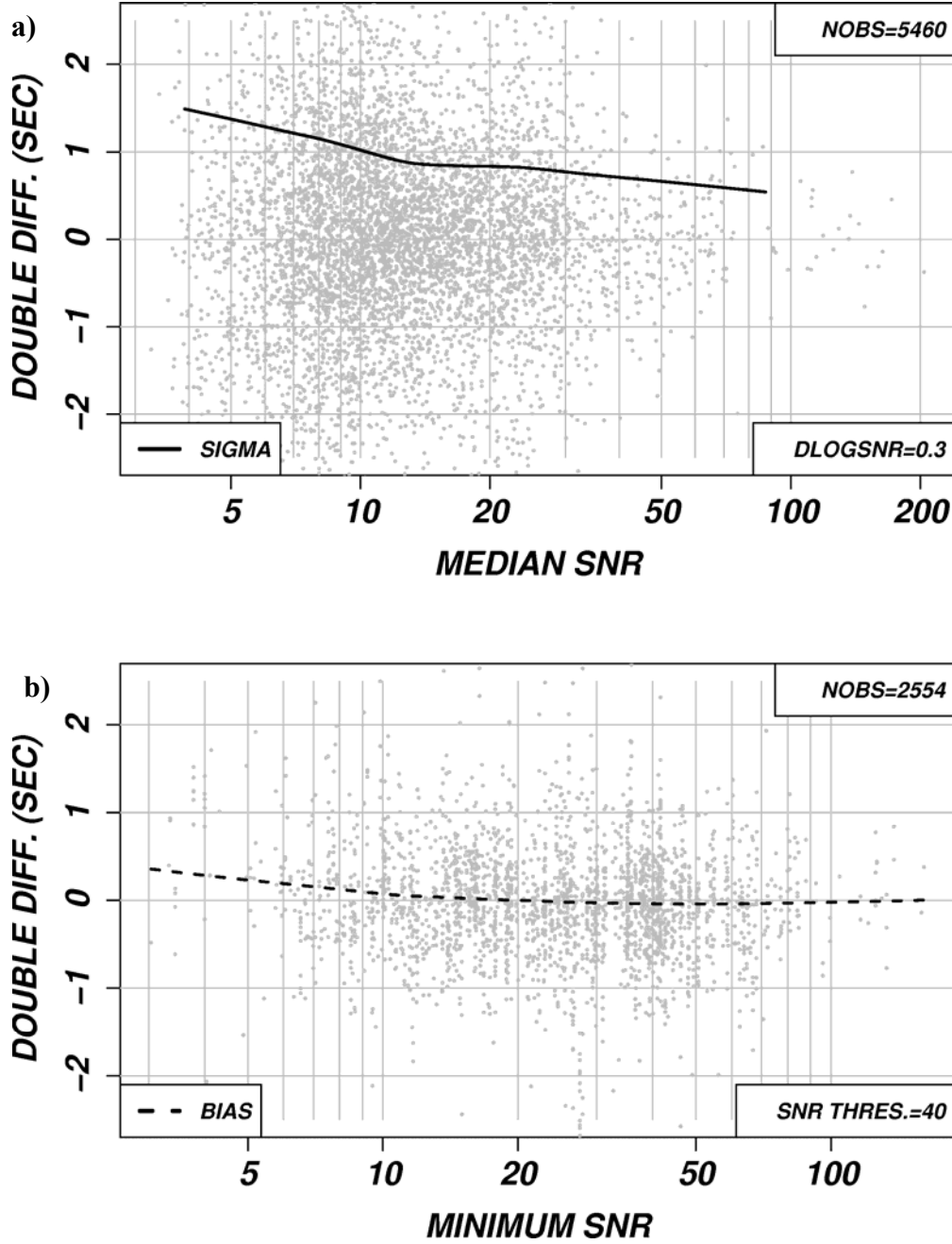


Figure 12. a) Double difference (DD) residuals where all four SNR are similar. The running SMAD (solid curve) represents the estimate of total variance. b) Double difference residuals where the SNR of three arrivals is large and the SNR of one arrival is small. The running median (dashed line) provides an estimate of the reading error bias.

Recall that the running SMAD in **Figure 12a** is an estimate of the total variance in the DD residuals which includes the background variance of travel-time prediction errors. We estimate Var_{hypo} from the second subset as the variance of DD residuals where the minimum SNR is larger than 40. This gives an estimate of $\sigma_{Hypo} = 0.24$ seconds which has to be removed from the total variance estimates in order to get an estimate of the reading error variance. **Figure 13** shows our resulting model for reading error bias and variance.

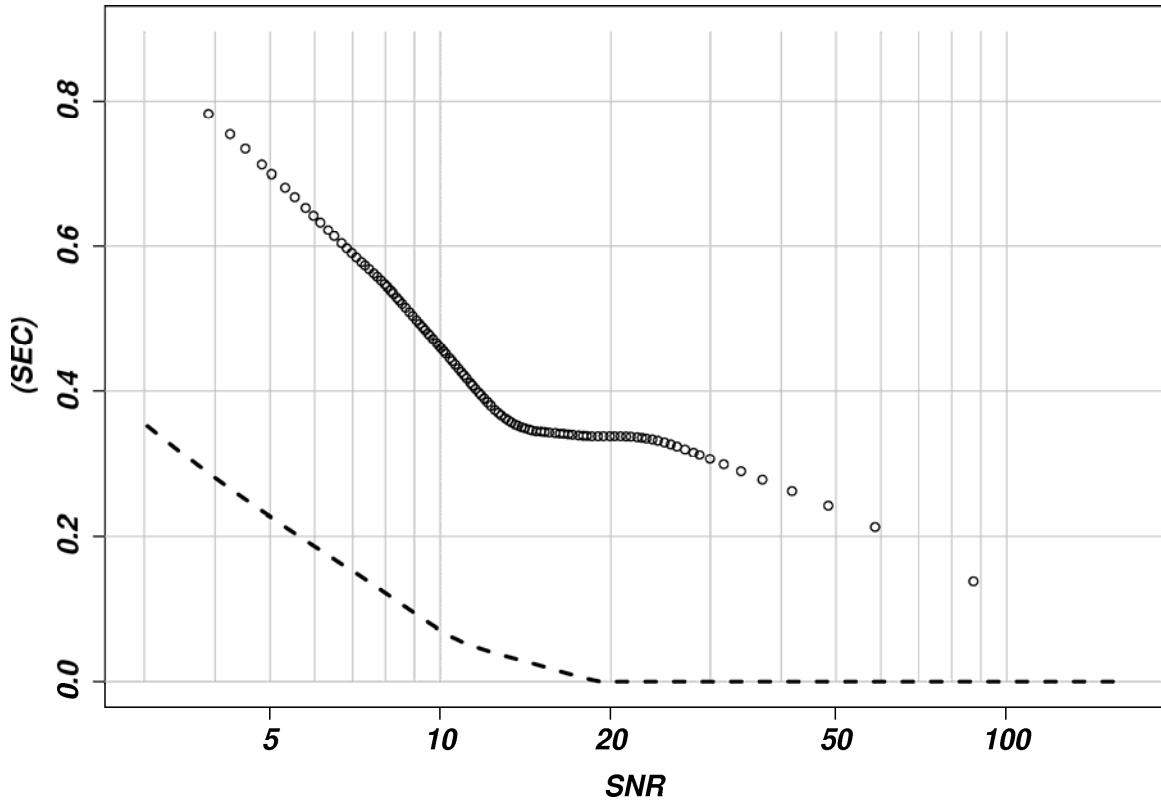


Figure 13. SNR-dependent model for pick delay (dashed line) and standard deviation (circles) derived from GT5 earthquakes with arrival and SNR data reported in the IDC REB.

The distributions of the P residuals of the GT5 clusters are shown in the histograms in **Figure 14** as corrections are progressively applied. The presentation of the histograms in **Figure 14** follows the format used for the Yucca Flat residuals in **Figure 10**. Gaussian density functions with mean and standard deviation of the empirical data are overlaid on the histograms. The histogram for the iasp91 residuals, r_{ij} , has virtually no bias (**Figure 14a**). The standard deviation is reduced when SSSCs are applied, but the residuals also become slightly offset (0.15 seconds) as shown in **Figure 14b**. **Figure 14c** shows that this offset is reduced by about 0.1 seconds when the bias of the measurement errors is corrected for using the estimated function of SNR. The histogram for the SSSC and bias corrected residuals normalized by standard deviations is shown in **Figure 14d**. The standard deviations were obtained as the square root of the sum of the variance of the measurement error model (**Figure 13**) and a flat value of 1 second, the latter representing the uncertainty in the SSSCs. The horizontal scale for these normalized residuals is in samples quantiles. The distribution of these normalized residuals deviates slightly from the standard Gaussian (zero mean, unit standard deviation) shown as a dashed line and it is somewhat skewed.

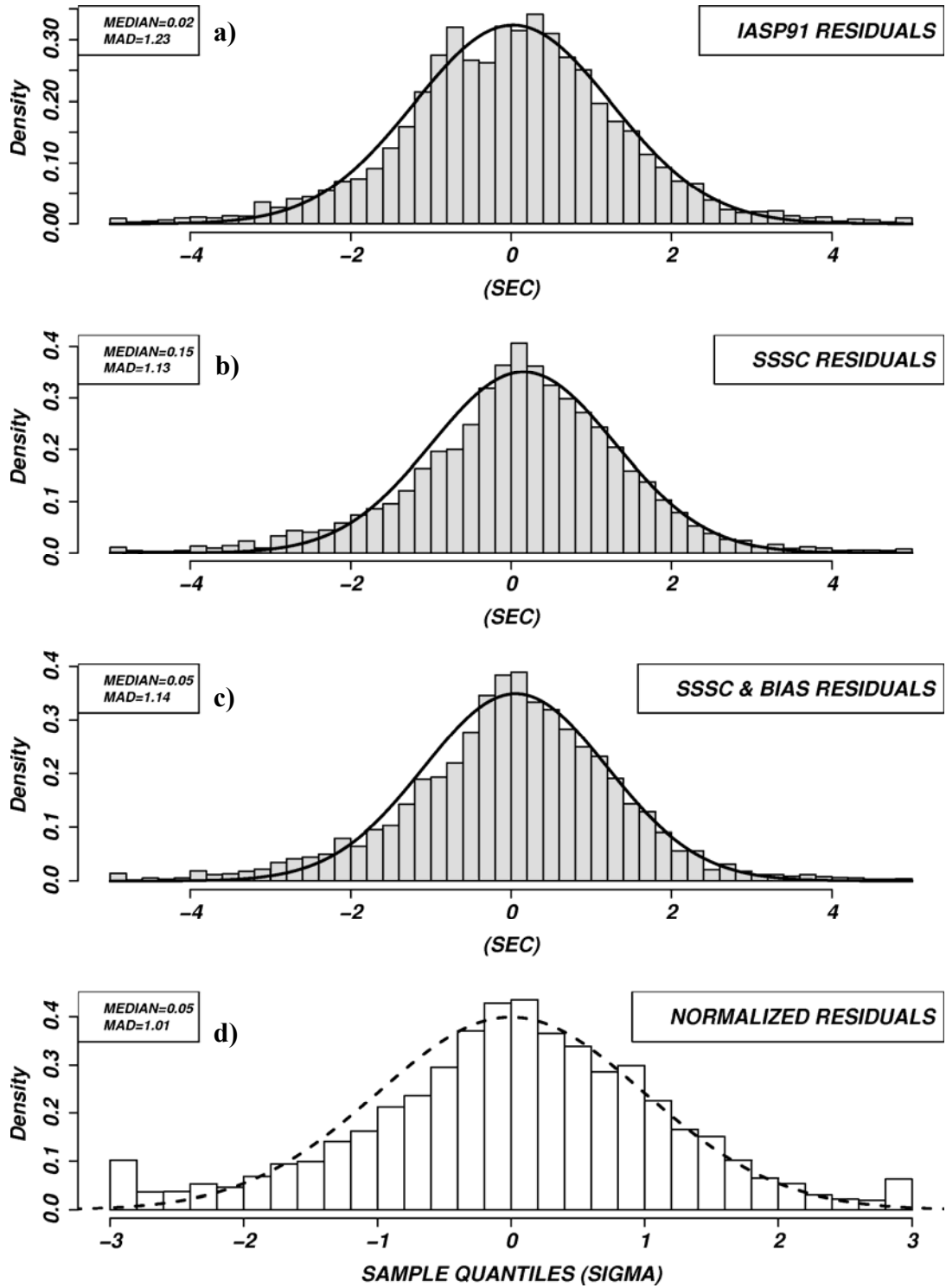


Figure 14. Histograms for arrival time residuals of GT5 events compared with Gaussian distributions. a) iasp91 residuals. b) residuals with SSSCs. c) residuals with SSSCs and SNR-based delay corrections. d) normalized residuals: path and bias corrected residuals are divided by the SNR-based standard deviation estimates. The horizontal scale for the normalized residuals is in sample quantiles and the density function of the standard Gaussian (zero mean and unit standard deviation) is drawn as a dashed line.

The distributions in **Figure 14** are fairly similar and with much less dramatic differences than the residual data of the Yucca Flat events in **Figure 10**, which can be attributed to the differences in ground truth for the two datasets. However, the histograms in **Figure 14** show some effect of variance reduction of the SSCs and of bias reduction from the measurement error corrections.

3.1.3. Location algorithm

3.1.3.1. Baseline location algorithm

Standard linearized location algorithms assume independent, zero-mean Gaussian errors and solve the inversion problem by an iteratively-weighted least-squares algorithm by minimizing the expression

$$(d - Gm)^T C^{-1} (d - Gm) \quad (10)$$

which is equivalent to solving the matrix equation

$$WGm = Wd \quad (11)$$

where G is the $(N \times M)$ design matrix containing the partial derivatives of N data by M model parameters, m is the $(M \times 1)$ model adjustment vector $[\Delta T, \Delta x, \Delta y, \Delta z]^T$, d is the $(N \times 1)$ vector of time residuals, C is the diagonal covariance matrix of *a priori* estimates of error variances (picking and model errors), and $W = C^{-1/2}$ is the diagonal $(N \times N)$ weight matrix. The $G_W m = d_W$ equation is often solved by singular value decomposition, which yields the general inverse

$$G_W^{-1} = V_W \Lambda_W^{-1} U_W^T \quad (12)$$

and thus the model adjustment of

$$m_{est} = G_W^{-1} d_W \quad (13)$$

At the $j+1^{\text{th}}$ iteration, the model vector is adjusted such that $m_{j+1} = m_j + m_{est}$.

Once a convergent solution is obtained, the location uncertainty is defined by the *a posteriori* model covariance matrix,

$$C_M = G^{-1} C G^{-1T} = V_W \Lambda_W^{-2} V_W^T \quad (14)$$

The model covariance matrix yields the two-dimensional error ellipse and one-dimensional errors for depth and origin time. The error ellipse encompassing the confidence region at a given α percentile level is defined by

$$(r - r_{loc})^T C_M (r - r_{loc}) = \kappa_\alpha^2 \quad (15)$$

where r_{loc} denotes the location vector of the epicenter. Jordan and Sverdrup (1981) recommended κ_α^2 as

$$\kappa_\alpha^2 = Ms^2 F_\alpha(M, K + N - M) \quad (16)$$

where the variance scaling factor s^2 is defined as

$$s^2 = \frac{K + \frac{1}{N} \sum d_w^2}{K + N - M} \quad (17)$$

and F_α is an F statistic with M and $K+N-M$ degrees of freedom at the critical level α with $M = 2$ and N observations. For $K = 0$, the scaling only includes the *a posteriori* data residuals (Flinn, 1965). Evernden (1969) pointed out that for small number of observations Flinn's approach yields unrealistically large "confidence" ellipses. Evernden (1969) assumed that the data uncertainties are known *a priori*, in which case κ_α^2 reduces to an α percentile χ^2 distribution with 2 degrees of freedom. Evernden's approach corresponds to $K = \infty$ and yields "coverage" ellipses. Thus, the choice of K offers a balance between "confidence" and "coverage" ellipses. Our implementation follows the IDC REB practice by setting K to a large value so that the formal uncertainty estimates approach "coverage" error ellipses. This linearized location algorithm constitutes our baseline algorithm, against which we measure improvements in location uncertainty estimates.

3.1.3.2. Improved location algorithm using non-diagonal data covariance matrix

In the presence of correlated systematic errors the data covariance matrix is no longer diagonal. The correlation structure implies that linear combinations of residuals may exist. This can be taken into account by diagonalizing the covariance matrix, thus reducing the dimensionality of the problem. The estimated location error ellipses then necessarily become larger, reflecting the fact that the number of independent data is less than the number of observations. Accordingly, our linearized location algorithm seeks a transformed set of equations $WGm = Wd$, in which the data covariance matrix is diagonal. To assure this, we solve the inversion problem in the eigen coordinate system in which the transformed observations are independent. Recall that the full data covariance matrix is the sum of network covariance and measurement error covariance matrices.

Our locator implements the SNR (or SNR surrogate) dependent systematic delays as a travel-time correction, while the SNR-dependent variances form the measurement error covariance matrix. Since we assume that the random picking errors are independent, their covariance matrix is diagonal. In other words, to obtain the full data covariance matrix, the measurement error variances simply add to the diagonal of the network covariance matrix.

The singular value decomposition of the full data covariance matrix is written as

$$C_D = U_D \Lambda_D V_D^T \quad (18)$$

where Λ_D is the diagonal matrix of eigenvalues and the columns of U_D contain the eigenvectors of C_D . We keep the first p largest eigenvalues from the cumulative eigenvalue spectrum such that 95% of the total variance is explained:

$$\sum_j^p \lambda_j / \sum_i^N \lambda_i \geq 0.95 \quad (19)$$

We then define p as the effective number of degrees of freedom of the data, with an $N-p$ dimension null space. Note that the 95% total variance level is somewhat arbitrary, but a conservative and workable choice.

Let $C_D = BB^T$, with $B = U_p \Lambda_p^{1/2}$, then the projection matrix

$$W = B^{-1} = \Lambda_p^{-1/2} U_p^T \quad (20)$$

orthogonalizes the data set and projects redundant observations into the null space. After applying the projections

$$G_W = \Lambda_p^{-1/2} U_p^T G \quad (21)$$

$$d_W = \Lambda_p^{-1/2} U_p^T d \quad (22)$$

the formalism remains the same as in the baseline algorithm, but now d_W represents linear combinations of the observed residuals, the “eigen residuals”. The major advantage of this approach is that the projection matrix is calculated only once for each event location.

The process is illustrated for a GT0 underground nuclear explosion carried out on 1992/03/26 at Pahute Mesa, Nevada Test Site. Pn was reported at 97 stations between 3° and 10° , the safest regional distance range for location purposes that excludes both the Pg/Pn crossover distance range and the Pn triplication zone. **Figure 15a** indicates a dense, but rather unbalanced network, with dense local networks in the Los Angeles basin and around the Mt. St. Helens volcano. **Figure 15b** shows the variogram constructed from the GT residuals for this event, overlaid by the variogram model obtained from the entire set of Pahute Mesa events (dark blue line) and our generic Pn variogram (red line). In this case the generic variogram model somewhat underestimates the correlation strength between nearby stations.

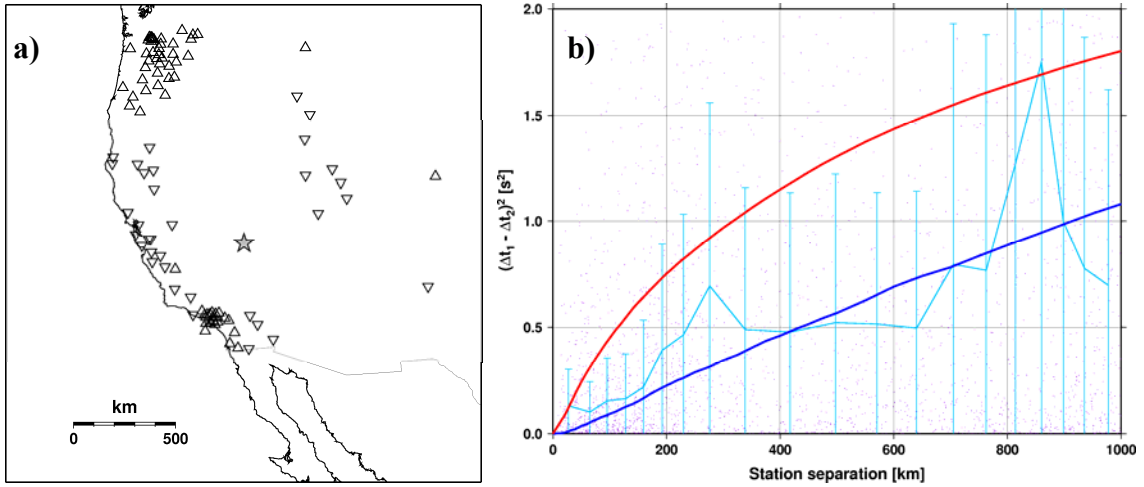


Figure 15. Regional network (3° - 10°) for the 1992/03/26 Pahute Mesa, NTS explosion (star). Triangles denote stations with positive Pn residuals, inverted triangles show stations with negative Pn residuals. Despite the small azimuthal gap (60°), the network is heavily unbalanced, owing to dense local networks in the Los Angeles basin to the Southwest and around the Mt. St. Helens volcano to the Northwest. b) Median residual difference squares in 5-percentile bins calculated from the GT residuals of the 1992/03/25 Pahute Mesa event (light blue line). The dark blue line shows the variogram model derived from the entire Pahute Mesa nuclear explosion cluster, the red line represents the generic Pn variogram model.

Figure 16a and **16b** show the network correlation matrix generated from the generic Pn variogram model as well as the corresponding cumulative eigenvalue spectrum. The network correlation matrix has been arranged by its nearest neighbor ordering of stations, and exhibits a quasi block-diagonal structure. Because many observations are strongly correlated, the first 40 largest eigenvalues explain 95% of the total covariance. In other

words, some 60% of the data are redundant. When the diagonal measurement error covariance matrix (assuming uniform, independent 1 second standard deviation reading errors) is added to the network covariance matrix (**Figure 16c and 16d**), the reading errors weaken the correlation strength, but leave the correlation pattern unchanged. Because of this “blurred” correlation structure, more eigenvalues are needed to explain 95% of the cumulative variance.

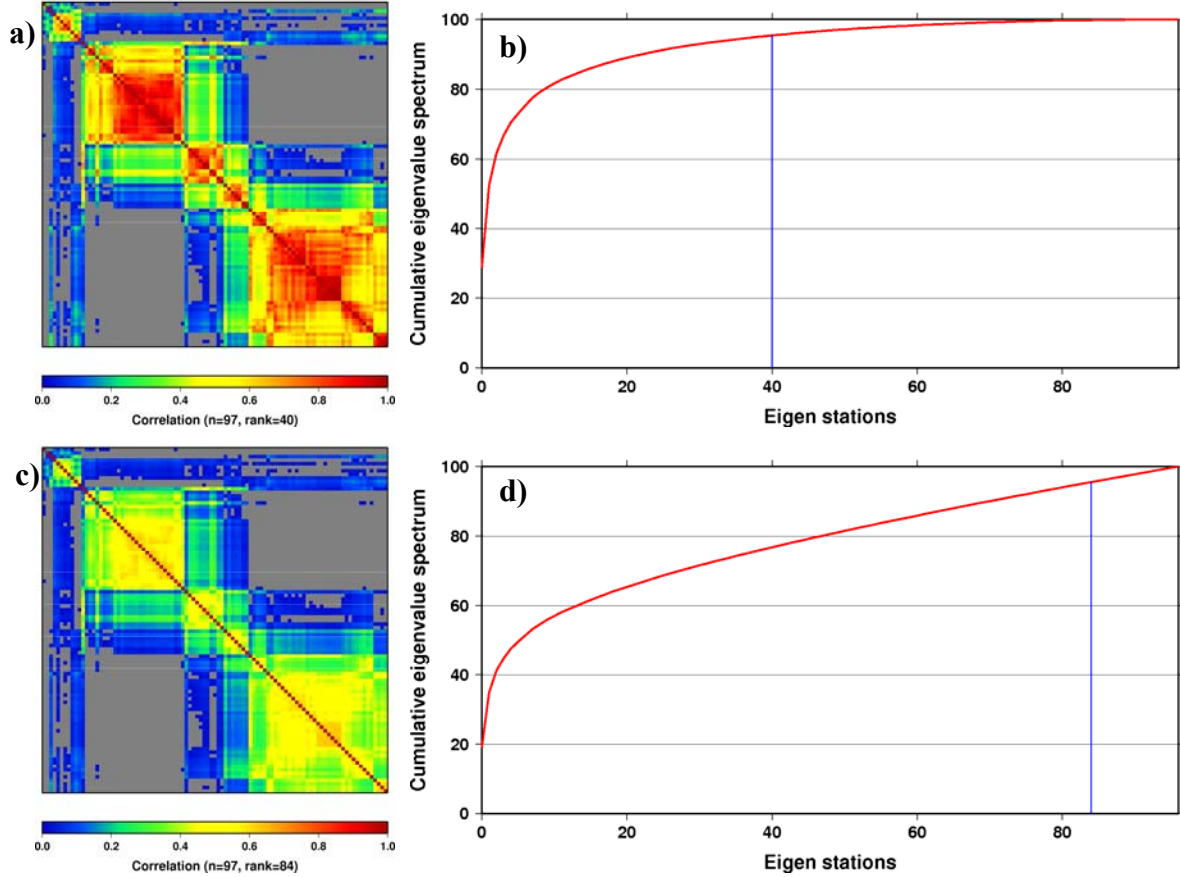


Figure 16. a) Station-station correlation matrix calculated from the regional network in Figure 5 and the Pn generic variogram model in Figure 4. The correlation matrix is arranged by the nearest-neighbor order of stations and exhibits a quasi block-diagonal structure. b) Cumulative eigenvalue spectrum of the network covariance matrix. The 40 largest eigenvalues explain 95% of the information carried by the 97-station network. c) Full data correlation matrix (with variances of reading errors added to the diagonal of the network covariance matrix). Reading errors weaken the correlation structure carried by the network. d) Because of the somewhat diluted correlation structure, more eigenvalues (83) are required to explain 95% of the covariance structure.

In general, when the picks suffer from large measurement errors, the correlation structure is blurred by the random noise of picking errors. Conversely, when the onsets are picked very accurately, the correlation structure becomes very important to obtain reliable location uncertainty estimates.

Figure 17 shows location results with the assumption of independent observations using iasp91 and CUB2 travel-time predictions, as well as the location obtained when correlated errors are accounted for. The calibrated travel-times reduce the 15 km iasp91

mislocation to 8 km. However, neither the uncorrelated iasp91 nor the uncorrelated CUB2 90% error ellipses cover the true location; they do not even overlap. When the correlation structure is accounted for, the location remains virtually the same as the CUB2 location, but the error ellipse now covers the GT location. Since the projected “eigenstations” are more uniformly distributed the shape of the error ellipse also becomes more circular.

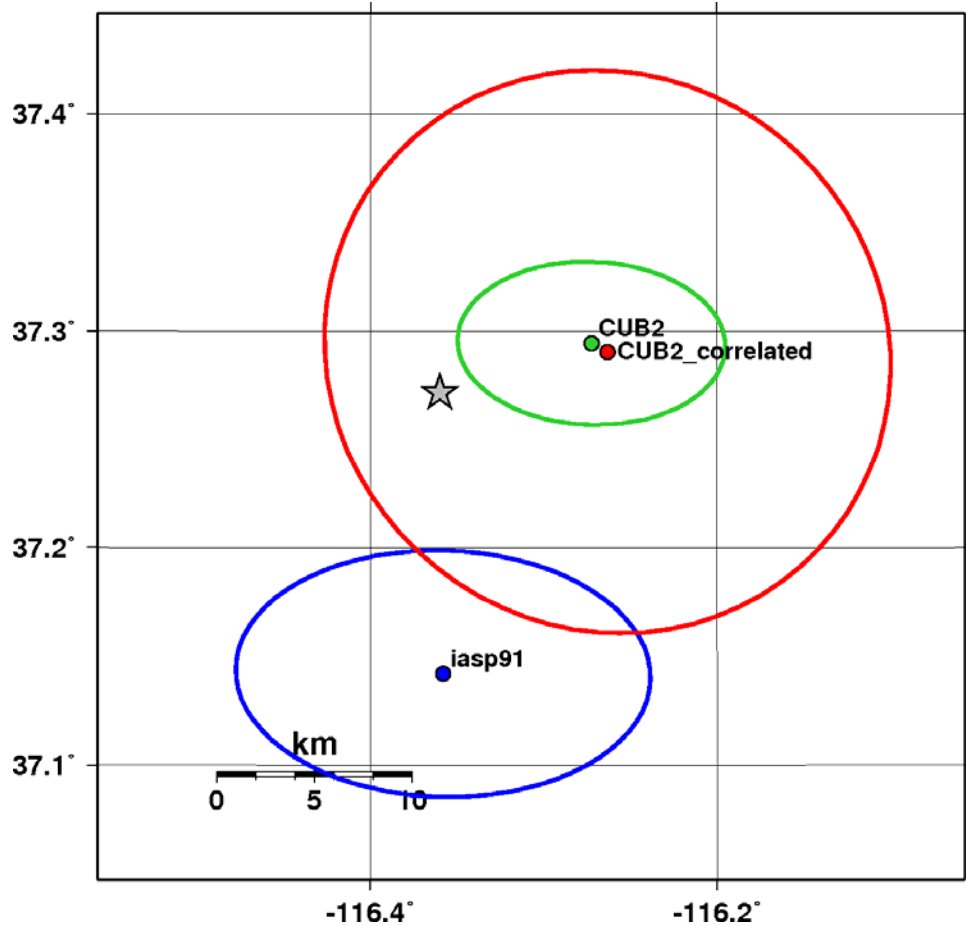


Figure 17. Relocation of the 1992/03/26 Pahute Mesa GT0 explosion (star) with iasp91 (blue) and CUB2 travel-time predictions assuming independent errors (green), and using CUB2 travel-times and the full data covariance matrix (red). When correlated errors are taken into account, the error ellipse covers the true location (star). The “eigenstations” represent a more balanced network resulting in a more circular error ellipse.

4. VALIDATION TESTS

We have performed a series of validation tests to demonstrate improvements due to SNR (or SNR surrogate) dependent phase pick delay and variance estimates and the representation of correlated model error structure. In order to assess the effects of improved estimates of measurement and model errors we first validate them separately, and then we investigate the combined effects of measurement and model errors.

4.1. Measurement errors

4.1.1. mb-distance dependent bias and variance

4.1.1.1. Relocation of 279 Yucca Flat explosions

To validate the mb-distance dependent errors we performed a location experiment, in which 279 Yucca Flat events were located with four different location algorithm configurations (**Table 2**) with regard to path-corrections, delay-corrections, and data weighting (SD-weighting). Data weighting refers to weighting with *a priori* standard deviations of reading and model errors. Estimated standard deviations of the path corrections were used as standard deviations of the model errors. Only first arriving P times were used and iasp91 served as a baseline travel-time model. To compare the four configurations we used four metrics: median bias, median error ellipse area, dimensionless standard error of the data fit (SDOBS), and percent coverage of error ellipses, i.e. the percentage of true locations covered by the 90% confidence ellipse.

Table 2. Location algorithm configurations for locating 279 Yucca Flats events.

Configuration	1	2	3	4
Path correction	No	Yes	Yes	Yes
Delay correction	No	No	Yes	Yes
SD-weighting	No	No	No	Yes

A priori errors were thus used only in configuration 4, for which estimated standard deviations of path corrections were used as model errors. The resulting metrics are listed in **Table 3**. Most of the bias of the standard location (*Configuration 1*) is removed by path-corrections (*Configuration 2*), but delay corrections (*Configurations 3 and 4*) do provide a small additional reduction in bias. *Configuration 2* provides a 52% reduction in bias w.r.t *Configuration 1*; *Configuration 3* further reduces the *Configuration 2* bias by 26%; and *Configuration 4* brings another 20% bias reduction w.r.t. *Configuration 3*.

Table 3. Location metrics versus location algorithm configurations of Table 1.

Configuration	1	2	3	4
Bias (km)	8.0	3.8	2.8	2.2
Error ellipse (km ²)	249	58	47	166
SDOBS (dimensionless)	1.2	0.54	0.54	0.56
Coverage (%)	70	60	88	95

The comparatively large reduction in location bias due to path corrections is expected as the path corrections vary between -1.5 and 4.0 seconds whereas the delay corrections are much smaller. Error ellipse coverage improves with delay corrections (*Configurations 3 and 4*). There is, however, excess coverage when SD weighting is applied (*Configuration 4*), suggesting that the *a priori* standard deviations are too conservative (too large). Delay corrections do not further reduce the fitting errors (SDOBS), which are

drastically reduced by path corrections. The drop in error ellipse area between *Configuration 1* and *Configuration 2* is consistent with the reduction in fitting error.

4.1.1.2. *mb* dependence for other test sites

There are several suites of widely reported underground nuclear explosions with GT1-2 ground truth locations at testing grounds other than NTS. The arrival-time data were, however, rarely recorded at regional distances, where the effects of measurement errors are most pronounced. Furthermore, there are no GT0 ground truth origin times and most of the suites have a limited range in magnitude due to testing practices. Although the testing program of a country may span a fairly large magnitude range, small explosions were often carried out at one site and large ones at another site. Therefore we analyze only data for Degelen Mountains of the Semipalatinsk testing grounds which includes explosions ranging from *mb* 4.4 to 6.

Since origin times are not known precisely, we turn to differences between arrival-time residual differences of station pairs, $r_{ij} - r_{kj}$, to estimate delay and variance functions. Such differences eliminate uncertain origin times. It is further assumed that one of the stations, k , generally records signals with large SNR so that delay and variance of the readings are negligible compared with those of other stations. Then differences between residuals at station, i , and those of the reference station, k , represent approximately the reading errors at the station, i , and the difference in path corrections between the two stations.

The Semipalatinsk testing grounds is known to be a bright spot for stations in Fennoscandia, which recorded large amplitudes with low dominant periods and high SNRs. Station HFS (Hagfors, Sweden) is used as a reference here as its readings were obtained from analog recordings with high time resolution (20 mm/s). The scatter diagrams in **Figure 18** compare residual differences between HFS and six other stations as a function of network *mb* for Degelen explosions. The delay (dashed lines) and standard deviation (dotted lines) functions are determined the same way as described in section 3.1.2.2. Delays as a function of *mb* are apparent although the magnitude range is about 1 magnitude smaller for Degelen than for NTS (1.5 compared with 2.5 magnitude units).

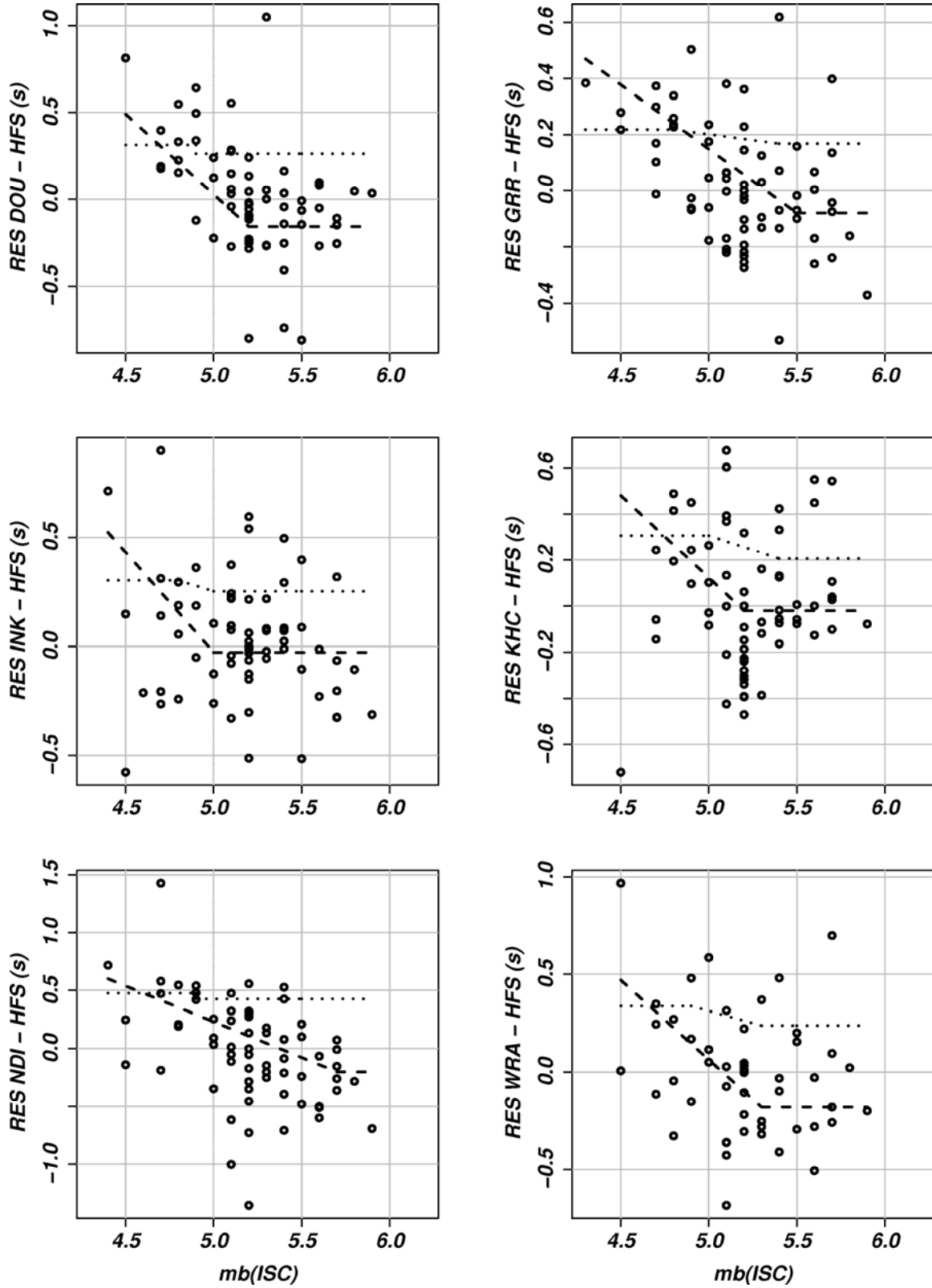


Figure 18. Differences in arrival time residuals for six stations (DOU, GRR, INK, KHC, NDI, WRA) relative to station HFS as a function of network $mb(ISC)$. Estimated bias and standard deviations as a function of $mb(ISC)$ are plotted as dashed and dotted lines.

Estimated delay and standard deviations for the differential residuals, $r_{ij} - r_{kj}$, are plotted in **Figure 19** for 41 stations. The delays and standard deviations are generally smaller than those obtained for NTS. For a given station the estimated delay represents the difference in delays between the station and that of HFS. The estimate therefore slightly under-estimates the actual delay. The estimated standard deviation for a given station, being the square root of the variance difference for the station and for HFS, slightly over-estimates the actual standard deviation. Hence, this analysis confirms that the model derived from NTS data applies to at least one other test site.

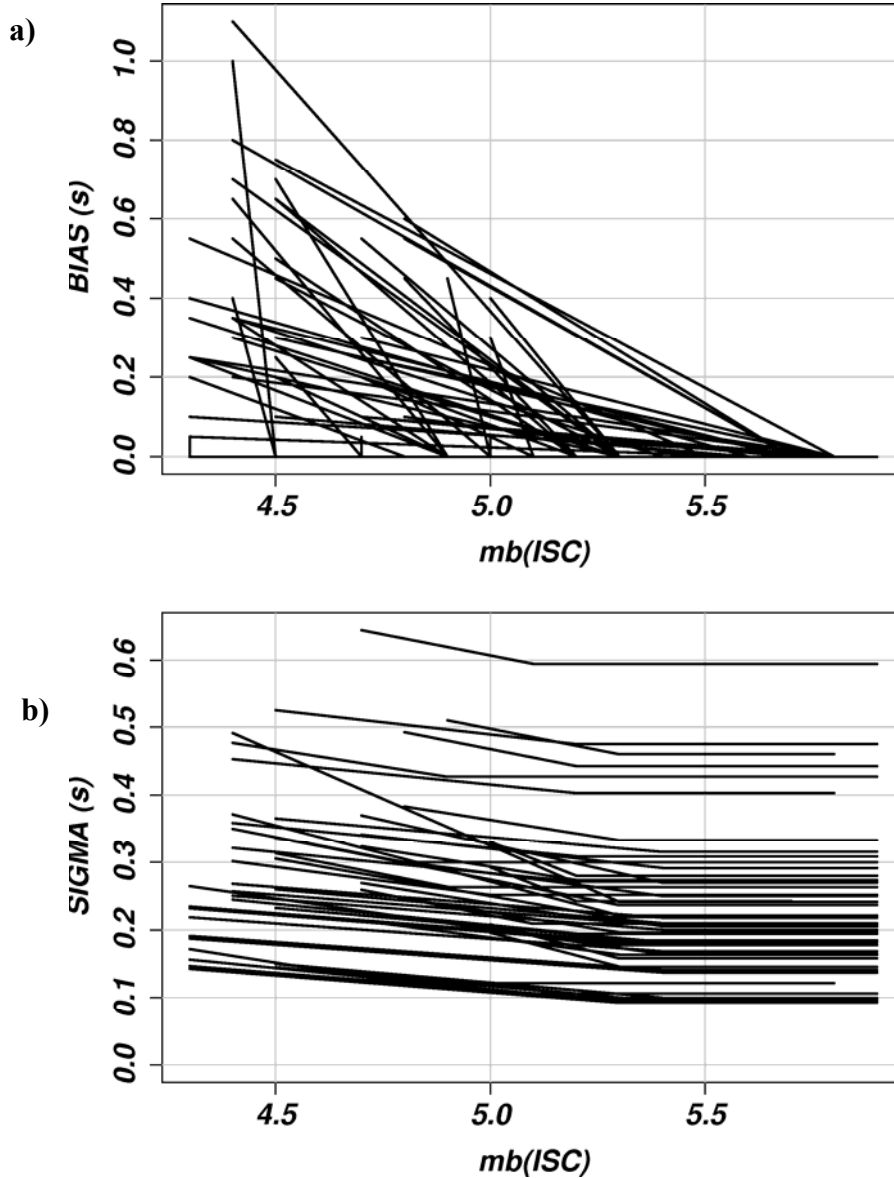


Figure 19. a) Estimated delay and b) standard deviation as a function of network mb(ISC) for Degelen explosions at 41 stations.

4.1.1.3. Transportability of the NTS mb-based measurement error model

The delay and standard deviation functions estimated with station pair differences for Degelen data above suggest that the error model obtained from NTS data are

conservative and probably over-estimates delay and standard deviation in reading errors for other areas and perhaps other event types. To assess transportability and general utility, the mb-error model was tested in several relocation experiments. **Figure 20** shows relocations of GT0-2 explosions in several areas with and without SSSCs and mb-delay corrections based on the NTS measurement error model using regional Pn only. The SSSCs were calculated from CUB2 (Shapiro and Ritzwoller, 2004) travel-time predictions relative to iasp91 (Kennett and Engdahl, 1991). The calibrated travel times are primarily responsible for location improvements, although they suffer from some remaining regional biases. The effect of mb-delay corrections is less obvious, but nonetheless significant. The amplitudes of the mb-delay corrections are much smaller than those from the SSSCs, therefore they only move events 1-2 kilometers maximum, which is often less than the GT accuracy. However, they do tend to tighten the event clusters containing a large range of magnitudes, by moving small events closer to large ones.

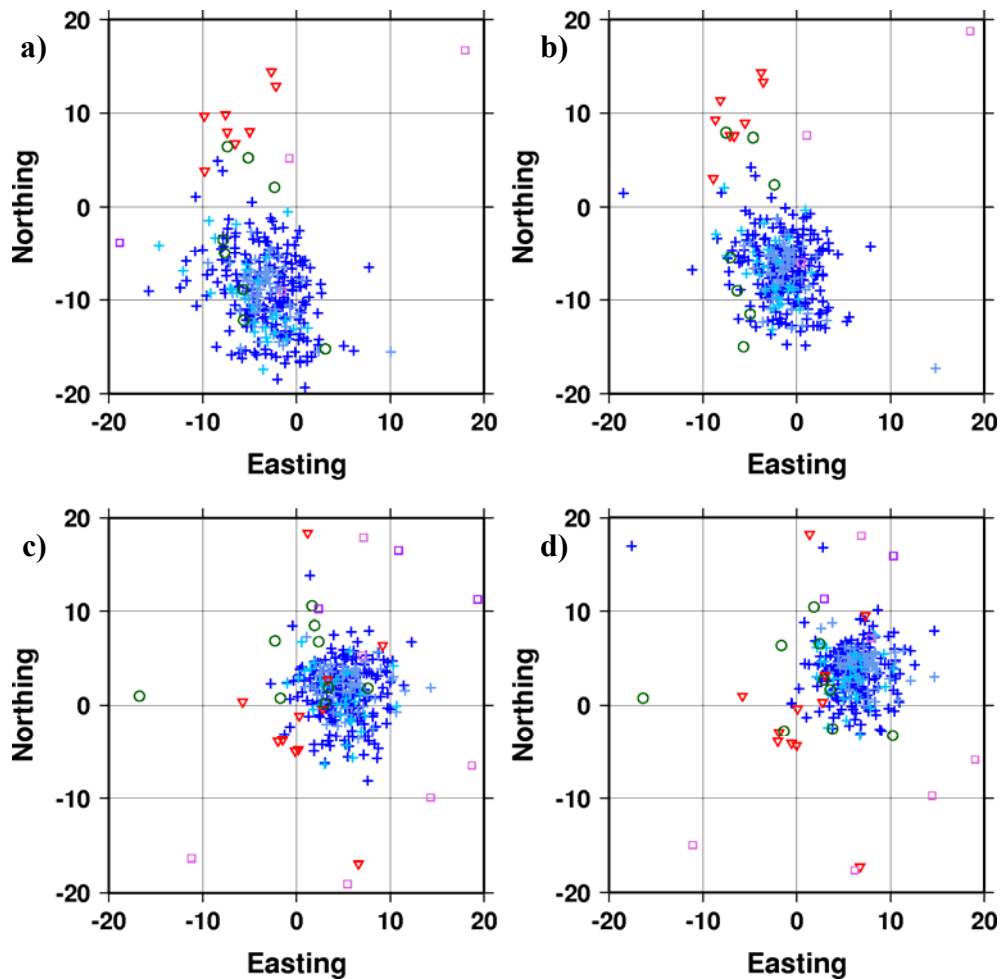


Figure 20. Mislocations (km) of NTS (crosses), Balapan (inverted triangles), Lop Nor (circles) and North Caspian (Azgir, Vega; squares) underground nuclear explosions using Pn arrivals with a) uncalibrated travel-time predictions (iasp91); b) uncalibrated travel-time predictions and mb-based delay and variance estimates c) calibrated (CUB2) travel-time predictions; d) calibrated travel-time predictions and mb-based delay and variance estimates. The mb-based delay corrections and measurement error estimates tighten 5 out of 7 clusters as listed in Table 4.

Since the location perturbations due to the mb-delay corrections and measurement error weights are on the order of the GT accuracies, a measure of the location improvements alone would be inevitably inconclusive. Instead, we measure the scatter of epicenters in a cluster around their center by calculating the covariance matrix of the point cloud described by the Easting and Northing coordinates of the mislocation measured from the ground truth location. Hence, the center of the point cloud represents the location bias of an event cluster, and the covariance matrix of the x-y coordinates around the cluster centroid characterizes the tightness of the cluster – the smaller the area of the ellipse defined by the covariance matrix, the tighter the cluster. **Table 4** summarizes the results for the Yucca Flat, Pahute Mesa, Rainier Mesa (NTS), Balapan (Semipalatinsk), Azgir and Vega (North Caspian PNEs) and Lop Nor (shaft area) underground nuclear explosions clusters. The numbers in parentheses represent the number of events in each cluster, located by at least 6 regional Pn phases. The area of the ellipse encompassing the point cloud at the 90% confidence level decreases for all clusters when SSSCs are applied, and further decreases for 5 out of 7 seven clusters when both SSSCs and mb-delay corrections are applied.

Table 4. Cluster metrics due to SSSCs and mb-delay corrections and weighting scheme for events with 6 or more regional Pn readings.

		Bias (km)	Area (km ²)
Yucca Flat, GT0 (201)	iasp91	9.5	231
	SSSC	5.2	153
	SSSC + mb	6.8	120
Pahute Mesa, GT0 (65)	iasp91	9.5	105
	SSSC	5.7	78
	SSSC + mb	8.3	58
Rainier Mesa, GT0 (29)	iasp91	10.8	193
	SSSC	4.7	124
	SSSC + mb	6.4	91
Balapan, GT1-2 (13)	iasp91	16.2	3001
	SSSC	1.2	980
	SSSC + mb	0.4	953
Azgir, GT1-2 (9)	iasp91	52.9	4726
	SSSC	24.7	970
	SSSC + mb	25.7	943
Vega, GT1-2 (13)	iasp91	49.0	6202
	SSSC	10.3	5698
	SSSC + mb	9.8	5759
Lop Nor, GT1-2 (9)	iasp91	7.6	898
	SSSC	2.7	485
	SSSC + mb	2.9	522

The mb-based error models were also tested with relocation experiments using over 2,000 globally distributed GT5 events (primarily earthquakes) produced by Bondár et al., (2007, 2008). We used both regional Pn and teleseismic P arrivals to relocate the GT5 earthquakes. Calibrated travel-time corrections (SSSCs) were calculated from the CUB2

(Shapiro and Ritzwoller, 2004) and J368D28 (Antolik et al., 2003) global three-dimensional models for Pn and P, respectively. Because the expected improvements (1-2 km) due to the mb-based delay corrections and weighting scheme are well within the 5 km location accuracy of the GT5 earthquakes, metrics on location improvement alone are inconclusive. Furthermore, since most of the earthquake clusters were formed from aftershock sequences of relatively narrow magnitude range, the mb-delay corrections often apply almost uniformly to the entire cluster. Demonstrable effects attributed to the mb-based error models can thus only be expected for clusters with a reasonably large magnitude range.

Figure 21 shows two such examples for the Izmit, Turkey and Kileaua, Hawaii earthquake clusters with events in the 4.0 – 5.7 magnitude range. The ellipses encompassing the point clouds (Easting and Northing coordinates in km, relative to the true locations) at one and two sigma levels are shown. Note the fairly regular separation of large and small events when they are located with uncalibrated (iasp91) travel-time predictions (**Figures 21a** and **d**). As we observed before, calibrated travel-times provide the first-order location improvements (**Figures 21b** and **e**). The mb-based delay corrections and weights tighten the clusters by moving small (red) events closer to large (blue) events (**Figures 21c** and **f**).

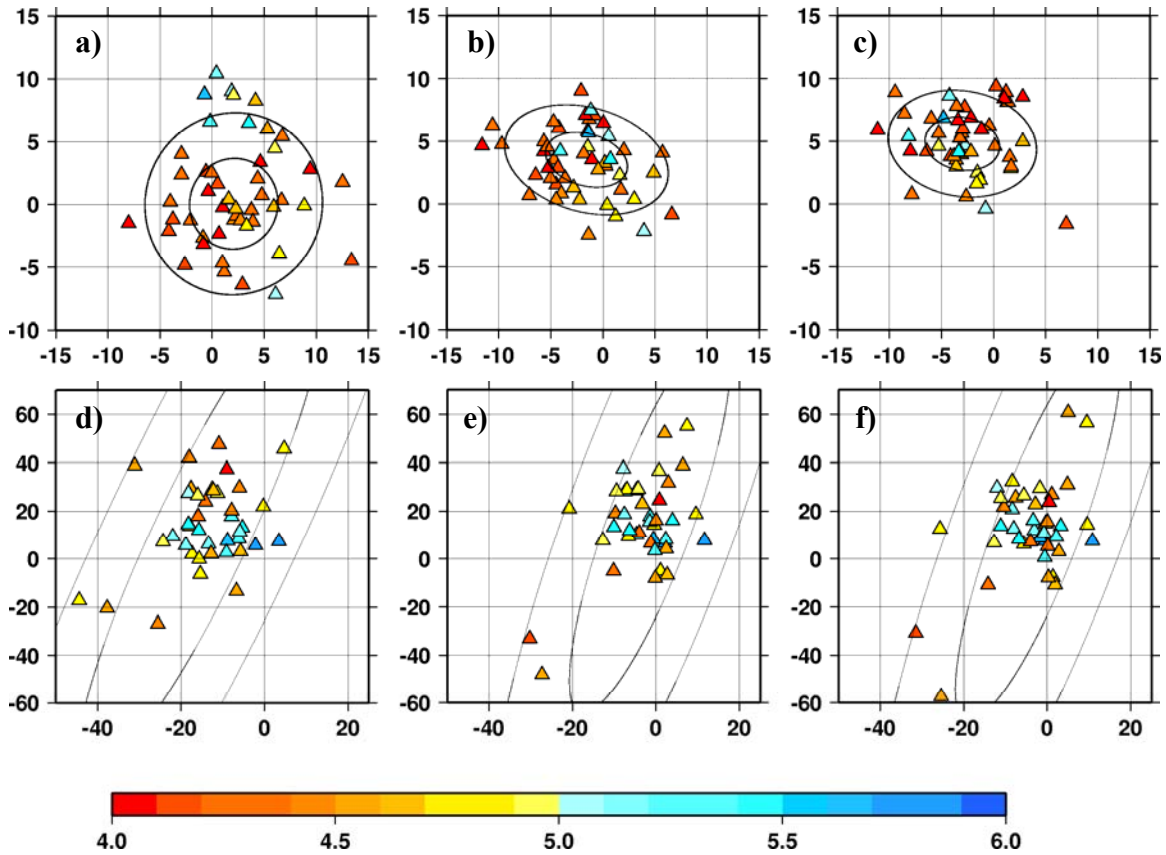


Figure 21. Mislocations of the Izmit, Turkey earthquake cluster using a) uncalibrated (iasp91); b) calibrated (CUB2, J362D28) travel-time predictions; c) calibrated travel-time predictions and mb-based delay and variance estimates. d-f are the same as a-c but for the Kileaua, Hawaii earthquake cluster. Events are color-coded by their mb and locations are plotted relative to the GT5 locations in

Easting-Northing coordinates (km). Error ellipses at the one and two sigma levels encompassing the point clouds are also shown. SSSCs are responsible for the bulk of location improvements; the mb-based delay corrections and weighting scheme further tighten the clusters.

For the Izmit cluster, the area of the ellipse defined by the covariance matrix of the point cloud and scaled to 90% confidence level decreases from 277 km² to 155 km² when calibrated travel-times are applied; mb-based delay corrections further decrease the ellipse area to 132 km². For the Kileaua cluster the trend is similar; SSSCs decrease the ellipse area from 10465 km² to 7100 km². The mb-based delay corrections and weights in this case slightly increase the ellipse area to 7269 km². However, it is visually clear that except for a few small events, the cluster gets tighter.

4.1.2. SNR-dependent bias and variance

Currently only the IDC REB routinely reports *sta/lta* SNR estimates for each arrival. Hence, any validation tests of SNR-dependent delay and variance models, derived from direct, *sta/lta* measures of SNR, must rely on events with station readings reported in the REB. This limited data availability restricted the location test to 493 GT5 earthquakes (Bondár et al., 2007) with 1,440 Pn and 5,057 P arrivals reported in the PIDC/IDC REB. These events are overwhelmingly dominated by teleseismic P for which the delay correction is quite modest. There are indications that the SNR-based delay correction tighten the clusters but the relocation tests are largely inconclusive and do not demonstrate significant improvement or degradation. A conclusive general utility test must await a future, much larger GT5 or more refined GT accuracy data set.

4.2. Correlated model errors

In the following tests we demonstrate improvements in coverage and location when correlated model error structure is taken into account. In these validation tests we use calibrated travel-time predictions from the CUB2 and J362D28 models for regional Pn and teleseismic P phases, respectively. To construct the non-diagonal data covariance matrices we employ our generic Pn and P variogram models to characterize the correlation structure in the model errors, and assume independent readings errors of 1 second standard deviation for both Pn and P picks. In all relocation experiments we fixed depth to the ground truth location. Arrivals were drawn from the EHB bulletin described above.

4.2.1. Coverage with increasing number of correlated stations

Figure 22 shows the mislocation, area of the 90% coverage error ellipse and the actual coverage with increasing number of stations for the 1994/10/07 Lop Nor and 1992/03/26 Pahute Mesa nuclear explosions we discussed earlier. For each teleseismic (Lop Nor) or regional (Pahute Mesa) subnetwork we selected those stations that provided the most uniform azimuthal coverage. Blue lines represent the solutions with the baseline location algorithm that ignores the correlated error structure; red lines show the results when correlated model errors are accounted for. When correlated errors are ignored, the error ellipses shrink indefinitely with increasing numbers of observations, losing coverage relatively early. On the other hand, when correlated errors are accounted for,

once the information content of the network is exhausted, the area of the error ellipse levels off and coverage is maintained across the entire range of subnetworks. Note that there are no significant improvements in location due to the fact that each subnetwork represents the most balanced network for the given number of stations. When correlated model errors are taken into account, the size of the error ellipse and the actual coverage no longer depend on the number of correlated stations used in the location.

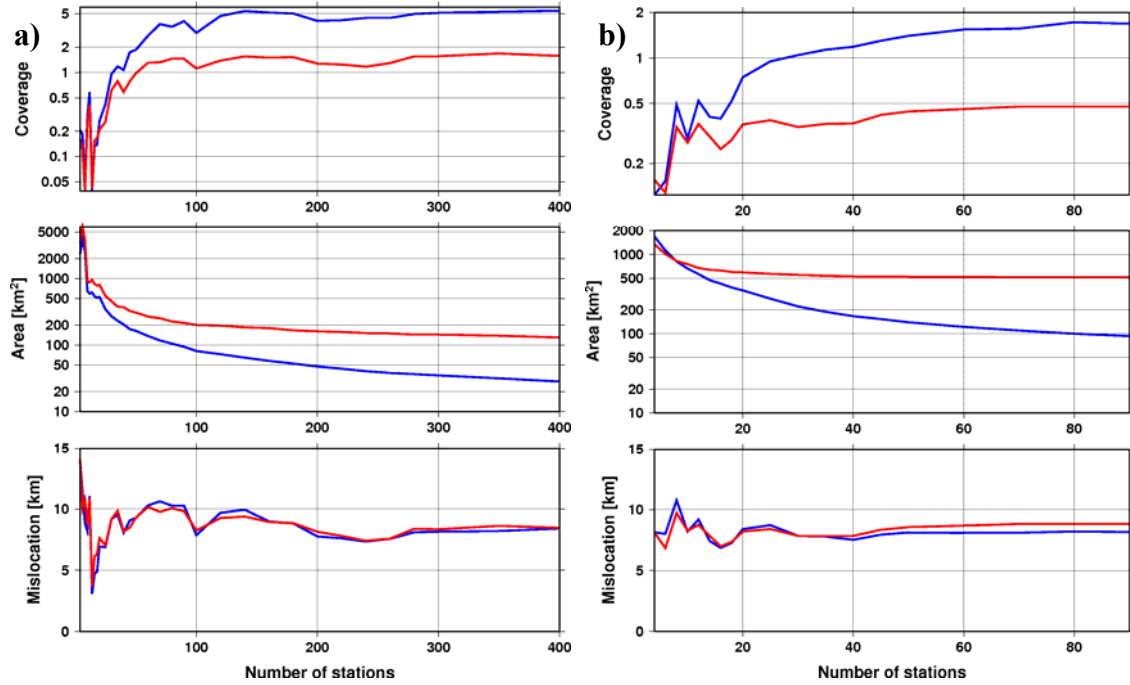


Figure 22. Mislocation (bottom), error ellipse area (middle) and coverage (top) with increasing number of optimally distributed stations when correlated errors are ignored (blue lines) and when correlated errors are accounted for (red lines) for the a) 1994/10/07, Lop Nor, and b) 1992/03/26, Pahute Mesa explosions. Coverage and error ellipse area are nearly independent of the number of stations for the correlated case.

4.2.2. Relocation of a large set of GT5 earthquakes

In our next validation experiment we relocated events from the GT5 earthquake clusters shown in **Figure 4**. For single-event location purposes this dataset of more than 2,000 events represents a wide variety of sparse and dense, balanced and unbalanced networks. We used both regional Pn and teleseismic P phases to locate events; in this case the full data covariance matrix becomes the composite of the Pn and P covariance matrices, assuming independence between regional (Pn) and teleseismic (P) stations.

Figure 23 shows the cumulative distributions of the mislocations, the semi-major axes of the 90% error ellipses and the coverage parameter when correlated errors are accounted for (red lines) and when they are ignored (blue lines). Because the mislocation improvements due to taking into account the correlated error structure are minor, the significant improvement in coverage is attributed to the increase in the length of the semi-major axis of the 90% error ellipse. Recall that the *a posteriori* model covariance matrix depends on the number of independent observations. Because the number of equivalent

uncorrelated data (linear combinations of residuals) is reduced, the formal location uncertainties described by the *a posteriori* model covariance matrix become larger, resulting in enlarged and more circular error ellipses. The fact that at higher percentiles the cumulative distributions of the semi-major axes are indistinguishable indicates that there are (most likely sparse) networks in the validation data set that do not suffer from correlated model errors. The enlarged, more realistic error ellipses provide better coverage. In fact, the actual coverage increases from 68% to 85% and approaches the theoretical χ^2 distribution with 2 degrees of freedom. The 90% error ellipses still do not cover the true locations 90% of the time, but the fact that we get much closer to the expected 90% coverage attests to the global applicability of our generic Pn and P variogram models.

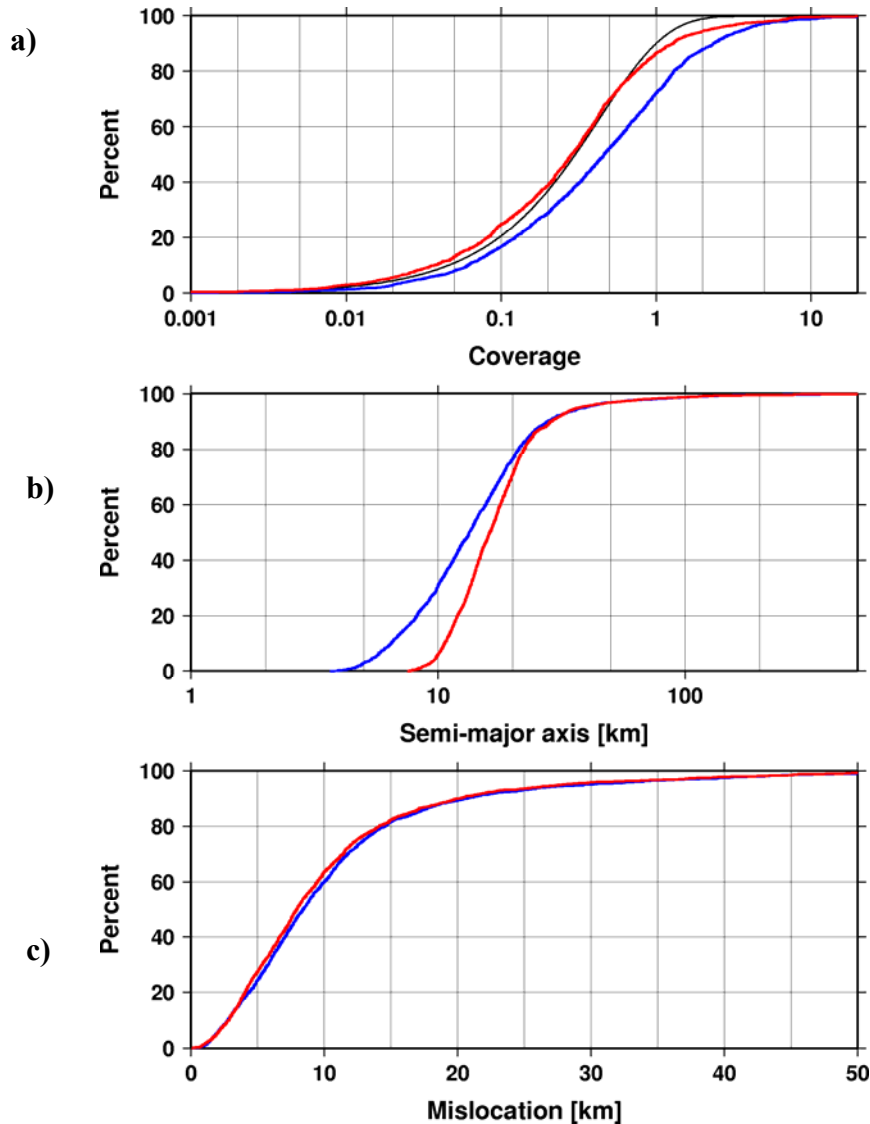


Figure 23. Cumulative distributions of GT5 earthquake a) coverage parameters, b) semi-major axes, and c) mislocations when correlated errors are ignored (blue) and when correlated errors are accounted for (red). Accounting for correlation increases actual coverage from about 68% to 85% and the cumulative more closely approaches the theoretical χ^2 distribution (black line in a).

We noted above that the location improvements are consistent, but minor. This is not surprising as one could expect significant location improvements only for heavily unbalanced networks where large numbers of correlated ray paths conspire to introduce location bias. **Figure 24** indicates that the first order effect of location improvements is attributed to the calibrated travel-time predictions and is negligible for events with magnitude less than mb 4.5.

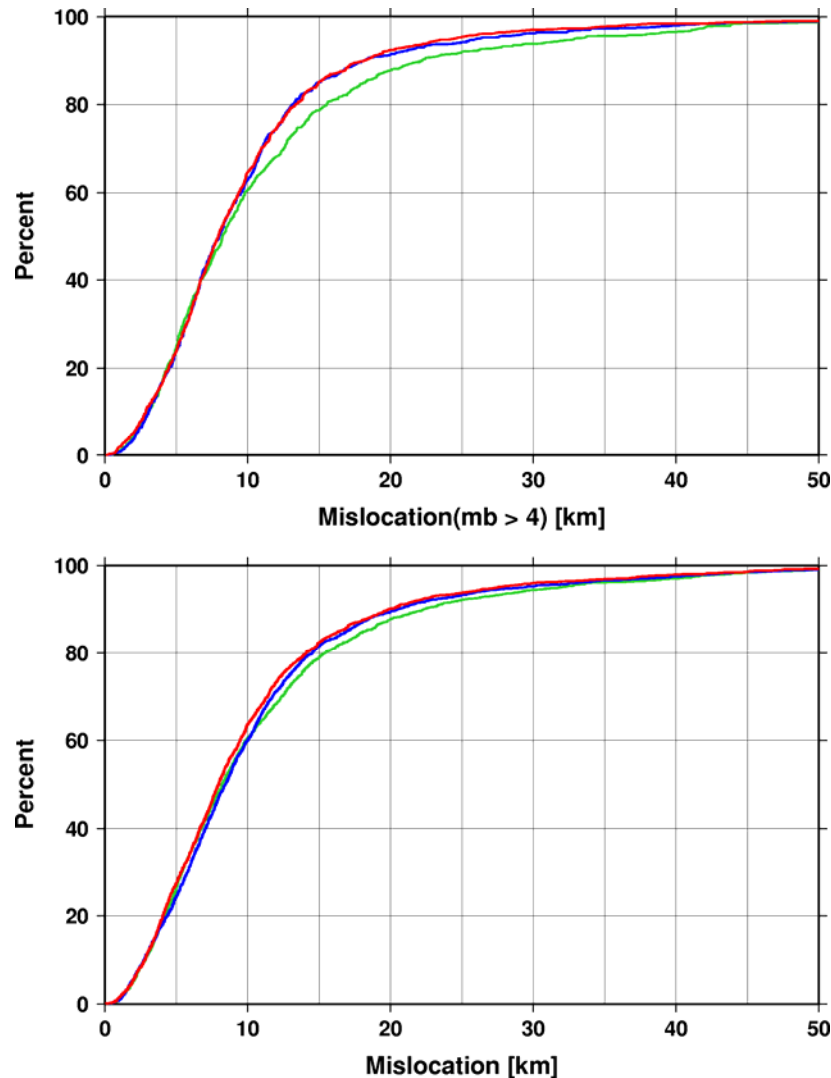


Figure 24. Cumulative distributions of mislocations of GT5 earthquakes with iasp91 (green), with calibrated CUB2 and/or J362D28 travel-time predictions (blue), and with calibrated travel-times with correlated errors (red). The upper panel shows the events with $mb \geq 4$. Calibrated travel-times are the first-order effects in location improvements; accounting for correlated model errors brings incremental location improvements.

It should be noted that both the CUB2 and J362D28 travel-time predictions are accompanied with isotropic, distance-dependent *a priori* model error estimates (Yang et al, 2004). One could argue that by simply increasing the *a priori* model errors one could achieve better coverage. Unfortunately, that approach does not work in general, as it

produces unrealistically large ellipses for networks with few stations, while the size of error ellipses still diminish with increasing number of stations. As we have shown earlier, one of the major advantages of incorporating correlated error structure in the location algorithm through the estimate of the full data covariance matrix is that it decouples the 90% error ellipse from the number of correlated observations. Indeed, when we plot the relocation results of the GT5 data set ordered by the number of observations used in the location (**Figure 25**), the median error ellipse areas are remarkably flat beyond 100 observations when correlated errors are incorporated into the location algorithm. Had we achieved 90% coverage, the 90th percentile curve would run along unity. Since we have only 85% coverage, the curve runs just above unity. Although sparse networks may also suffer from correlated model errors, it is more likely that dense networks have enough correlated ray paths to introduce location bias if the correlated error structure is ignored. Taking into account the correlated model error structure reduces this bias, hence the location improvement.

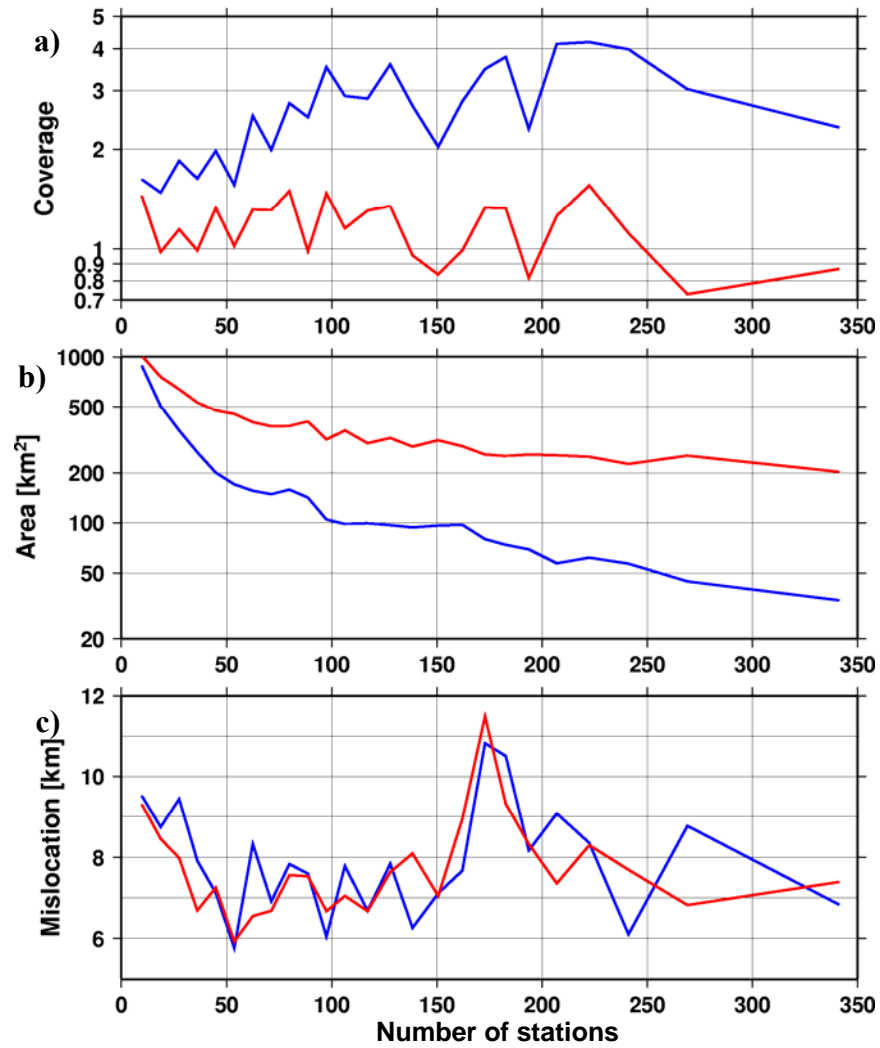


Figure 25. a) 90% coverage, b) median error ellipse area, and c) median mislocation of 2275 GT5 earthquakes ordered by the number of stations used in the location when correlated errors are accounted for (red), and when they are ignored (blue). Coverage and error ellipse area are less dependent upon the number of stations for the correlated case.

4.2.3. Sparse, unbalanced networks

As we noted above, sparse, unbalanced networks may also suffer from correlated errors. In order to demonstrate the consequences of ignoring correlated model errors, we consider the Matochkin Shar, Novaya Zemlya test site with 29 GT1-2 (Kohl et al., 2003) underground nuclear explosions. While these explosions were recorded by hundreds of teleseismic stations, the regional coverage is extremely poor. **Figure 26** shows the regional stations between 0° and 15° . Most of the stations are located in Fennoscandia, and the secondary azimuthal gap is invariably larger than 270° . The numbers in the parentheses show the number of events reported by a particular station. None of the explosions were recorded by more than 10 stations and only three events were reported by both KBS in Spitsbergen and NRI in Northern Russia.

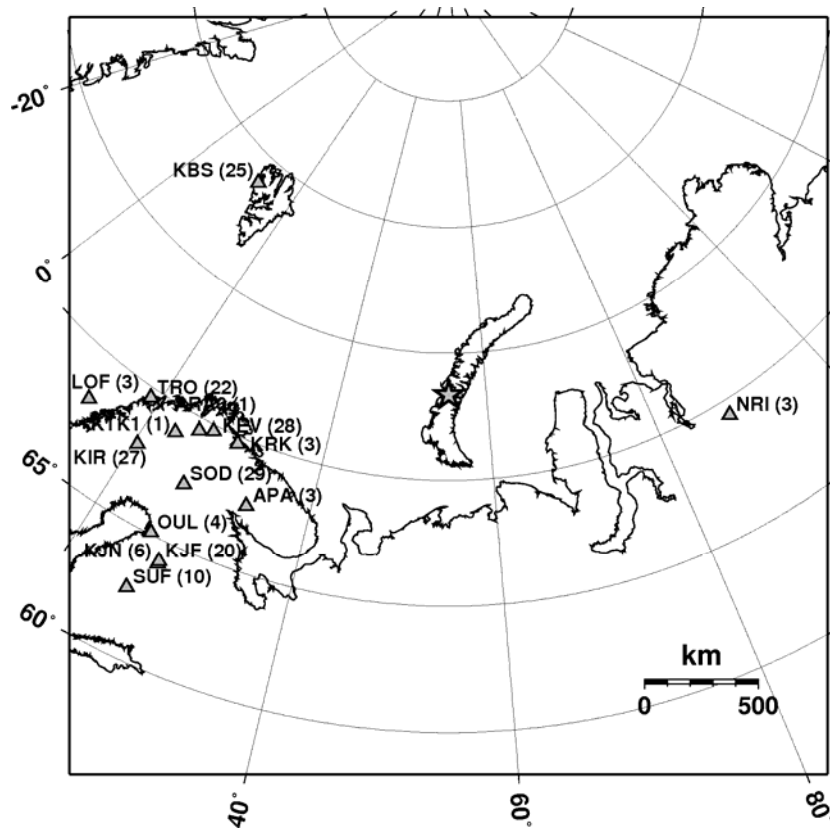


Figure 26. Regional network (0° - 15°) for the Matochkin Shar, Novaya Zemlya test site (star). The numbers in the parentheses indicate the number of events (out of 29 GT1-2 explosions) reported by a particular station. Most of the stations are located in Fennoscandia, representing a sparse, heavily unbalanced network.

Figure 27 shows the relocation results using uncalibrated (iasp91) travel-times and assuming independent errors (green triangles), and using calibrated (CUB2) travel-time predictions and accounting for the correlated model error structure (red circles). The events are ordered by the number of reporting stations (shown at the right hand-side of the plots). Not surprisingly, it is difficult to obtain accurate locations with such an unbalanced network. The baseline location algorithm (1D travel-times and ignoring the

correlated error structure), produces huge error ellipses, and because of the basically flat misfit surface, trades off location with origin time. As before, calibrated travel-time predictions are responsible for the bulk of location improvements by reducing the median mislocation from 71 km to 53 km. Accounting for correlated errors brings further incremental location improvements, with a median mislocation of 51 km. However, the major effect manifests itself in the error ellipses, which get significantly smaller; calibrated travel-times alone reduce the median area of the error from 54,000 km² to 25,000 km². Taking into account the correlated error structure brings down the median error ellipse area to 18,000 km². Hence, analogously to the dense network end of the spectrum, the size of the error ellipse does not grow indefinitely with decreasing number of stations when correlated errors are accounted for.

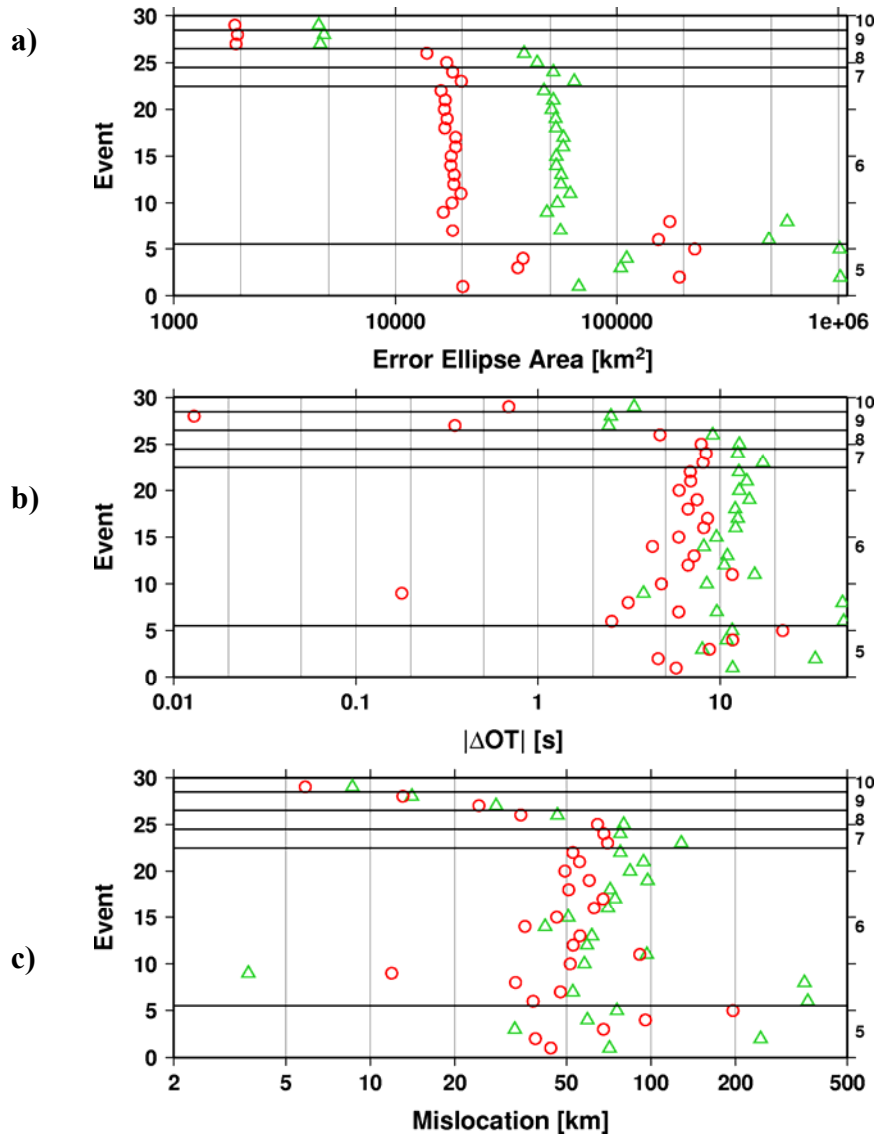


Figure 27. a) 90% error ellipse area, b) origin time difference, and c) mislocation for 29 GT1-2 Matochkin Shar nuclear explosions using iasp91 travel-times and assuming independent errors (green triangles), and using calibrated (CUB2) travel-times and accounting for the correlated model

error structure (red circles). The horizontal bands represent the number of defining phases, indicated on the right hand axis.

Figure 28 shows the misfit surfaces obtained from grid searches for the 1975/08/23 8:59:59 Novaya Zemlya underground nuclear explosion. This event was recorded by five regional stations between 0° and 15° (KBS, KEV, KIR, SOD and KJF). The comparison of 90% confidence error ellipses with the misfit contours indicates that for such unfavorable station geometry the error ellipse provides a poor representation of the actual misfit when uncalibrated (iasp91) travel-time predictions are used assuming independent errors. In this case the linearization (ignoring higher order terms in the Taylor expansion) step, leading to the assumption of a quadratic misfit surface is hardly valid, especially at high confidence levels. Calibrated (CUB2) travel-times improve the linearization as the misfit surface becomes more quadratic. Finally, the misfit surface is the steepest and most elliptical around the global minimum when correlated errors are accounted for.

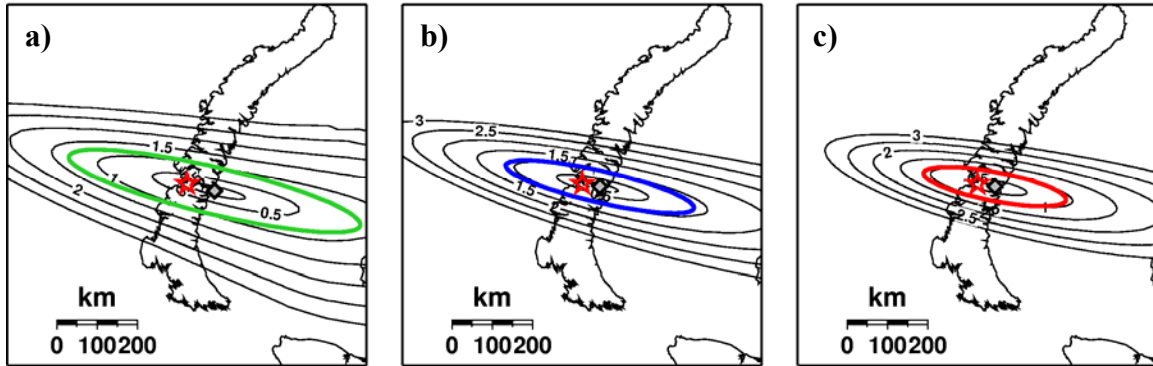


Figure 28. Grid search misfit surfaces for the 1975/08/23 8:59:59 Novaya Zemlya underground nuclear explosion (red star) using regional Pn from KBS, KEV, KIR, SOD and KJF. Misfit contours from the grid search (thin lines) and 90% confidence error ellipses from the iterative least squared relocations (thick, colored lines and grey diamonds) are shown when a) uncalibrated (iasp91) travel-times assuming independent errors, b) calibrated (CUB2) travel-times assuming independent errors, and c) calibrated travel-times with correlated model error structure were used. The misfit surface is the steepest when correlated errors are accounted for.

It should be noted that the error ellipse areas are still too large and invariably provide 100% coverage. This means that the generic Pn variogram model overestimates the correlated model errors for this region. The example we presented here represents a region where further improvements can be achieved by carrying out region/site-specific calibration, employing 3D velocity models or empirical travel-time corrections and then subsequently developing a refined model of the remaining correlated errors.

4.3. Correlated model error structure vs non-Gaussian errors

We have shown that incorporating correlated model errors into a linearized iterative least squares location algorithm yields more reliable location uncertainty estimates and location improvements for unbalanced networks. One might wonder how important is the correlated error structure compared to non-Gaussian errors and/or non-linear effects? In this section, we quantify non-Gaussian distribution of residuals using GT data and then

quantify their expected impact on the location algorithm with and without correlated errors.

Many researchers have pointed out that the distribution of travel-time residuals is non-Gaussian and onset times tend to be picked late with decreasing event size (e.g. Anderson, 1982; Billings et al., 1994; Buland, 1986; Douglas et al., 1997, 2005). Indeed, **Figure 29** shows that the skewed generalized extreme value (GEV) distribution provides a better fit to the distribution of GT residuals (residuals calculated with respect to the GT location using CUB2 and J362D28 travel-time predictions for Pn and P, respectively) than the symmetric Gaussian distribution. The GEV was first introduced by Jenkinson (1955) and it combines the Gumbel, Fréchet and Weibull extreme value distribution families. The GEV probability density function is defined by

$$f(x, \mu, \sigma, \xi) = \frac{1}{\sigma} \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi - 1} \exp \left(- \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right) \quad (23)$$

where μ is the location parameter, σ is the scale parameter, and ξ is the shape parameter which governs the tail behavior of the distribution. Note that except for the right tail, the shapes of the best fitting Gaussian and GEV distributions are quite similar, which probably explains why the Gaussian error assumption fares reasonably well for seismic event location.

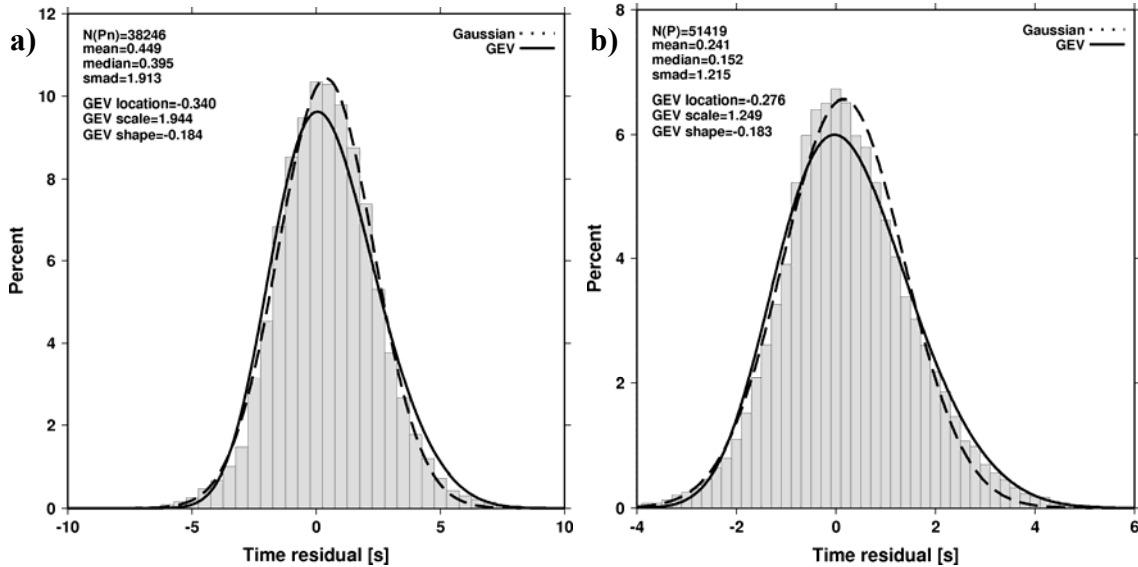


Figure 29. GT residual distributions of a) regional Pn and b) teleseismic P phases from the GT5 validation data set. Solid and dashed lines represent the best fitting Generalized Extreme Value (GEV) and Gaussian distributions, respectively.

We should emphasize that the linearized iterative least squares algorithm we used to incorporate the correlated model error structure was simply a choice of convenience, so that we could compare results with the baseline linearized iterative least squares algorithm. Our methodology does not call for any particular minimization algorithm. Once the data are projected into the eigensystem where the observations are independent, one could use iteratively-reweighted least squares using any l_p norm, a grid search, a

simulated annealing algorithm or any other non-linear location algorithm to solve the inversion problem (i.e. to minimize the objective function).

4.3.1. Dense networks

In order to quantify the relative importance of correlated error structure versus non-Gaussian errors we performed synthetic Monte Carlo experiments. In the first experiment we used the network recording the 1992/03/26 Pahute Mesa GT0 nuclear explosion again. We chose a real network that exemplifies real-world non-uniform spatial station distribution. The event was recorded by 97 regional stations between 3° and 10° (**Figure 30a**), and 183 teleseismic stations between 28° and 90° (**Figure 30b**). We constructed the corresponding *a priori* full data covariance matrices using the generic Pn and P variogram models. We then generated 5,000 realizations of independent Gaussian random variables for the 97-station regional network, as well as 5,000 realizations of Gaussian random variables with the correlation structure described by the full data covariance matrices for Pn. We repeated the same process for the 183-station teleseismic network. When generating the multidimensional independent, identically distributed (iid) and correlated, identically distributed (cid) random variables, for the marginal distributions we used the parameters of the best fitting Gaussian distributions (**Figure 29**) for Pn ($\mu = 0.395$, $\sigma = 1.913$) and P ($\mu = 0.152$, $\sigma = 1.215$). We also generated 5,000 realizations of iid and cid random variables using the best fitting GEV distributions for Pn ($\mu = -0.340$, $\sigma = 1.944$, $\zeta = -0.184$) and P ($\mu = -0.276$, $\sigma = 1.249$, $\zeta = -0.183$).

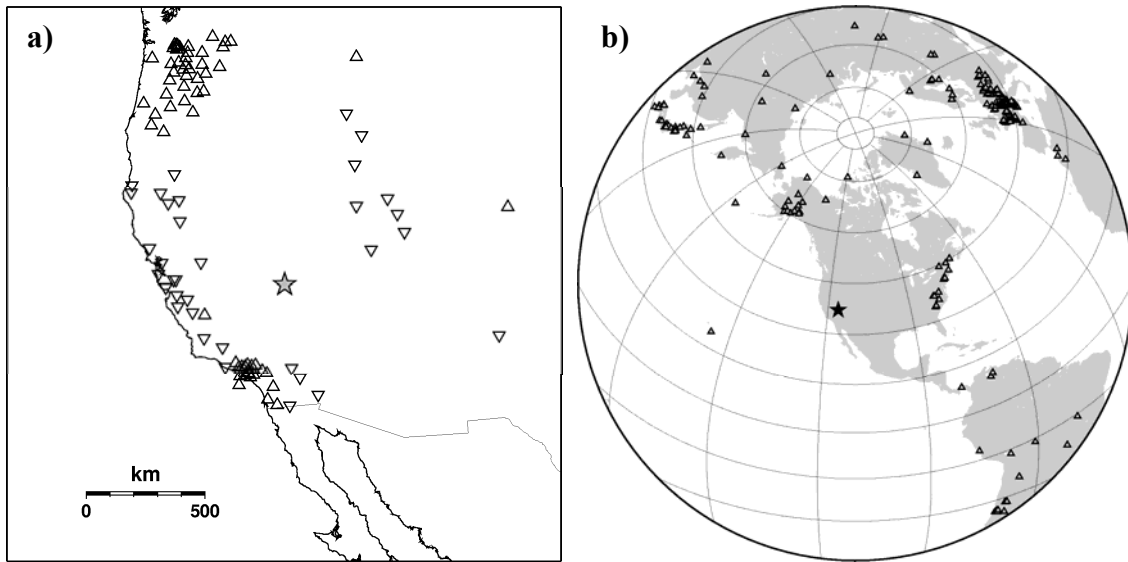


Figure 30. a) Regional (3° - 10°), and b) teleseismic network (28° - 90°) for the 1992/03/26 Pahute Mesa, NTS underground nuclear explosion (star). Both the regional and teleseismic networks are heavily unbalanced, with concentrations of stations in Cascadia and the LA basin (regional), and Europe, Japan and Alaska (teleseismic).

Generating correlated random variables is typically far from trivial when the joint distribution is not a multidimensional Gaussian distribution. One of the most attractive features of a copula formulation (see **Appendix**) is that it separates the dependence structure from the marginal distributions. Hence, copula theory allows an elegant and fast

method to generate multidimensional random variables with a prescribed dependence structure and arbitrary marginal distributions (Yan, 2006). The dependence structure itself is fully described by the copula. Copulas come in various flavors and families, each representing specific dependence structures. Linear correlation, represented by the Gaussian copula, is just one of the many possible dependence structures. In our case we specified the copula as a multidimensional Gaussian copula characterized by the full *a priori* data covariance matrix, and drew samples from Gaussian and GEV marginal distributions, respectively. The procedure for generating correlated random errors with Gaussian and GEV marginal distributions is described in the **Appendix**.

In summary, we generated eight separate sets of 5,000 realizations of multidimensional random variables: iid and cid Gaussian for P_n and P, as well as iid and cid GEV for P_n and P. The random variables represent total errors (that is, the combination of model and measurement errors) and were added to the predicted travel times at the corresponding stations. In each realization we located the event assuming that the errors are independent; and assuming that the errors are correlated. By specifying the dependence structure as independent or correlated allowed us to measure the effect of correlated errors on location uncertainties and mislocation, as well as the effect of making the wrong assumptions about the nature of the error structure. Because the correlation structure due to unmodeled velocity heterogeneities remains the same, generating random variables from GEV marginal distributions allowed us to measure the effect of non-Gaussian errors on location and coverage. The experiments are best summarized in a truth table, as shown in **Table 5**.

Table 5. Truth table for location algorithm assumption versus Monte Carlo synthetic errors.

	Gaussian errors		GEV errors	
Locator assumption	Independent	Correlated	Independent	Correlated
Independent errors	<i>True_iid</i>	<i>False_iid</i>	<i>True_iid</i>	<i>False_iid</i>
Correlated errors	<i>False_cid</i>	<i>True_cid</i>	<i>False_cid</i>	<i>True_cid</i>

True_iid denotes the cases when the actual errors were independent and we located the event correctly assuming that the errors were independent; *False_iid* stands for the cases when the errors were correlated, but we still located the event assuming independence; *False_cid* represents the cases when the errors were truly independent but we assumed that they were correlated; finally, *True_cid* denotes the cases when the errors were correlated and we rightly accounted for the correlation structure. Hence, *False_iid* and *False_cid* stand for the cases where we made the wrong assumption about the model error dependence structure.

Figure 31 summarizes the regional location results for the various cases we described above. Not surprisingly, the results are best when the right assumption is made about the error structure. The coverage statistic for the *True_iid* (dotted line) and *True_cid* (solid line) cases are virtually indistinguishable, providing an actual coverage very close to 90%. The coverage statistics are similar because the *True_cid* coverage no longer

depends on the number of correlated stations. The mislocations are larger in the presence of correlated error structure, simply because we now have fewer independent observations (84 vs 97). Interestingly enough, when we assume that the errors are correlated when they are in fact independent (*False_cid*, dashed line), we end up with overly conservative coverage, and even better locations than *True_cid*. Thus, forcing a correlated error structure on independent observations does no harm. On the other hand, ignoring an existing correlated structure in the data (*False_iid*, dashed-dotted line) results in abysmal (20%) coverage and the largest mislocations.

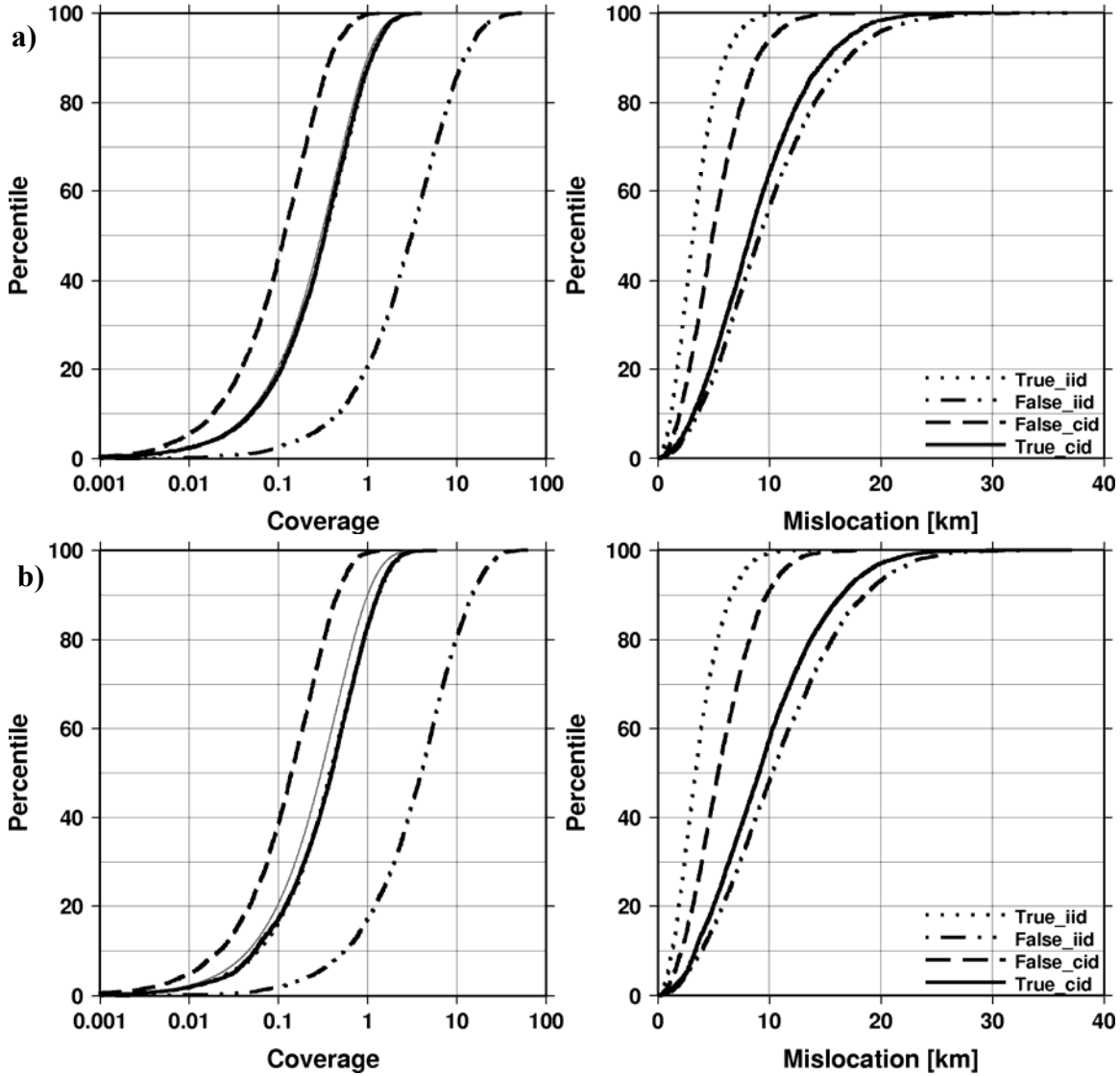


Figure 31. Regional Monte Carlo location experiment with the 1992/03/26 Pahute Mesa explosion with a) Gaussian, and b) GEV residuals. *True_iid* (dotted line): independence assumption with independent errors; *False_iid* (dashed-dotted line): independence assumption with correlated errors; *False_cid* (dashed line): correlated assumption with independent errors; *True_cid* (solid line): correlated assumption with correlated errors. The thin grey line represents the cumulative distribution of the coverage parameter for the expected 90th percentile χ^2 distribution with 2 degrees of freedom.

Figure 32 indicates that for the teleseismic network, the effect of right or wrong assumptions is similar to those we have seen in the regional case. Again, the penalty is the most severe for ignoring the correlation structure present in the data.

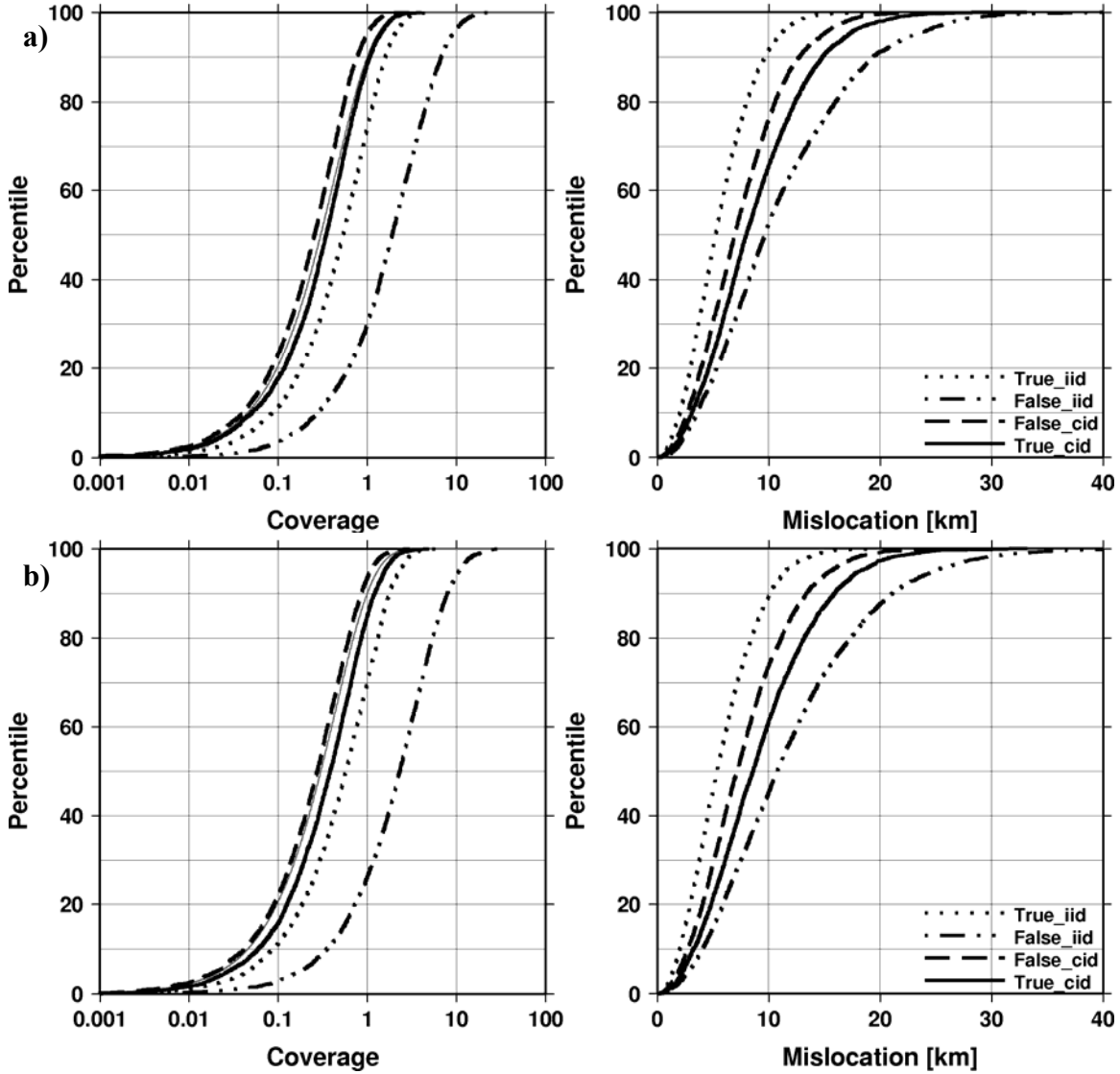


Figure 32. Teleseismic Monte Carlo location experiment with the 1992/03/26 Pahute Mesa nuclear explosion with a) Gaussian, and b) GEV residuals. The legend is the same as in Figure 31.

When the errors are not Gaussian, but drawn from the heavier-tailed GEV distribution, both the coverage and the location becomes somewhat worse (lower panels in **Figures 31** and **32**). However, albeit consistent, the deterioration is only minor. Indeed, **Figure 33** indicates that non-Gaussian errors make the coverage and location significantly worse only at the highest percentiles (that is, less than 5% of the cases). Thus, taking into account the correlated error structure is a first-order effect in improving location uncertainty estimates; heavy-tailed non-Gaussian error distributions are a second-order effect. It should be noted that ignoring the correlated error structure (*False_iid*) is affected the most by non-Gaussian errors.

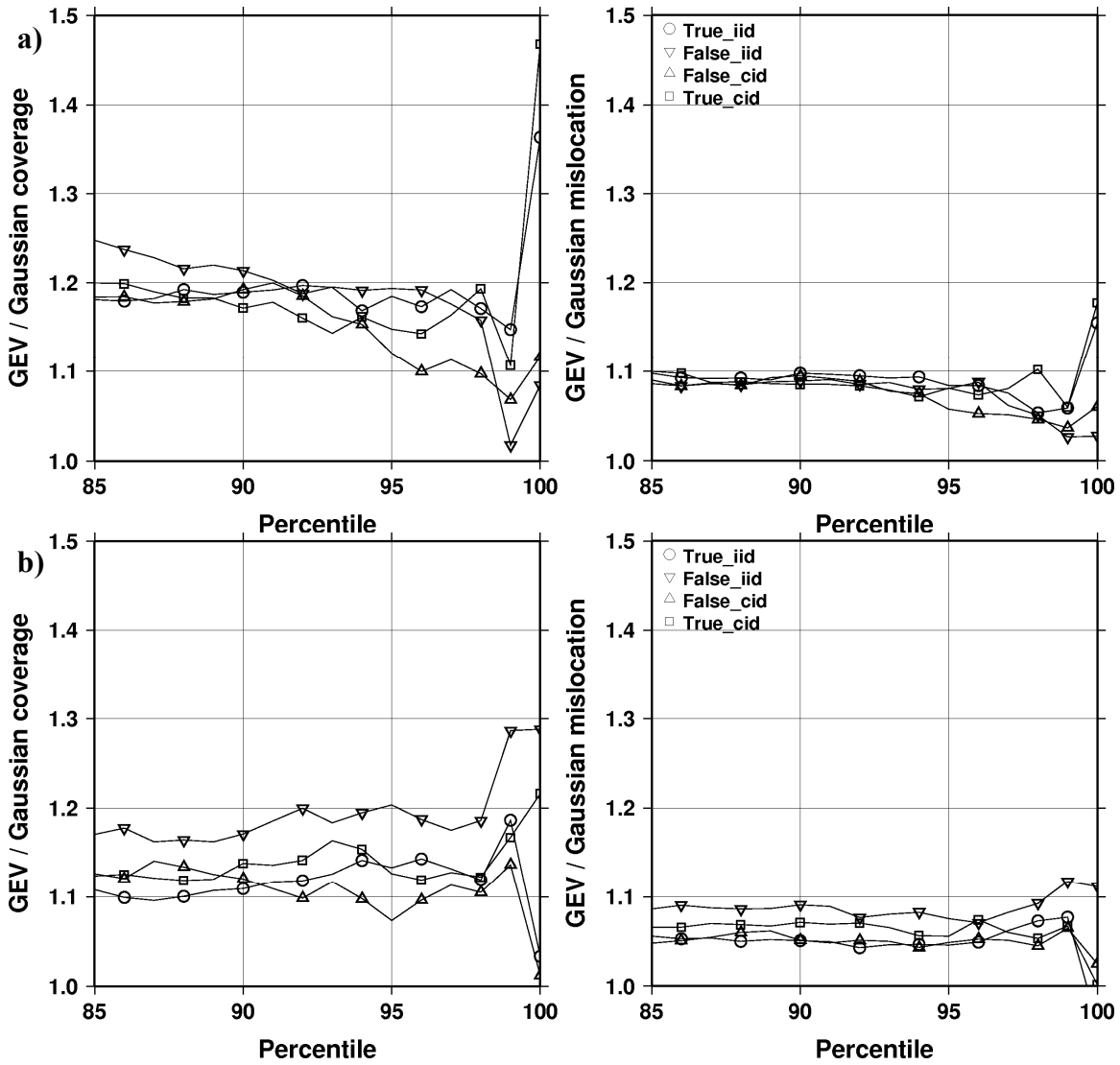


Figure 33. Ratios between coverage and mislocation results at each percentile when GEV and Gaussian error distributions were used as Monte Carlo inputs. a) Regional, b) teleseismic case. The differences due to non-Gaussian errors become significant at the highest percentile levels.

4.3.2. Increasingly correlated networks

Our next exercise was designed to quantify how coverage is affected with increasingly correlated networks when the correlated error structure is ignored. We chose the networks for three GT0 events with large numbers of regional Pn observations in the 3° and 10° distance range: the 1988/07/07 Pahute Mesa (53 Pn) and 1988/10/13 Yucca Flat (54 Pn) underground nuclear explosions, as well as the 1992/11/02 Switzerland ammunition storage explosion (57 Pn). **Figure 34** indicates that the Pahute Mesa and Yucca Flat explosions were not reported by all stations in the dense local networks in the Los Angeles basin and Cascadia, which made the regional network so unbalanced for the 1992/03/26 Pahute Mesa event we used in the previous validation test.

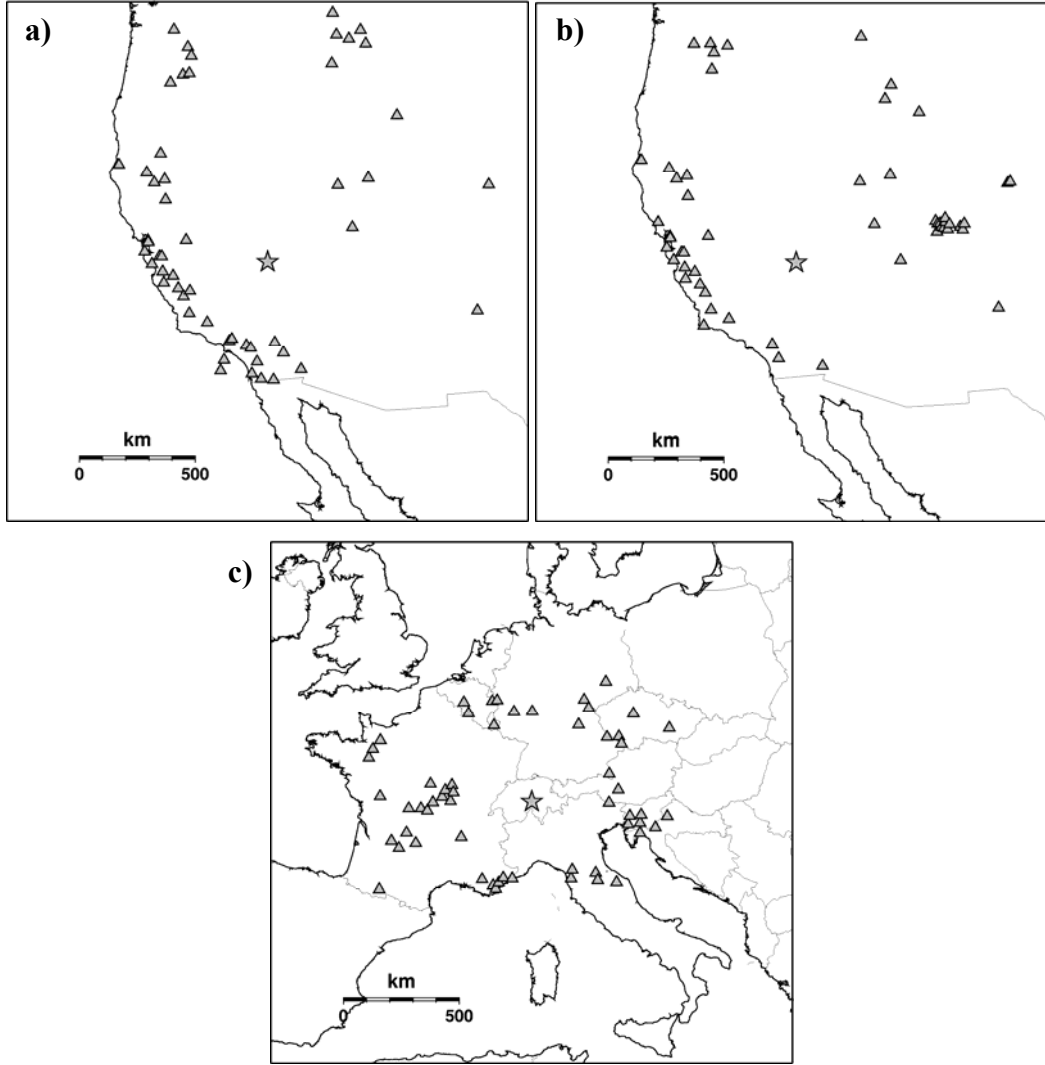


Figure 34. Regional networks (3° - 10°) for the a) 1988/07/07 Pahute Mesa, b) 1988/10/13 Yucca Flat, and c) 1992/11/02 Switzerland GT0 explosions.

We performed a synthetic Monte Carlo experiment by randomly selecting 10, 20, 30 and 40-station subnetworks (150 realizations per subnetwork), and as in the previous experiment, we located the events with the assumptions of independent and correlated errors, where the predicted arrival times were perturbed by random independent/correlated variables drawn from Gaussian/GEV marginal distributions.

Figure 35 summarizes the results. The procedure for generating random correlated errors from Gaussian and GEV marginal distributions is described in the **Appendix**. Open circles represent the *True_iid* case (the locator assumed that the errors were independent and they were indeed independent); inverted solid triangles indicate the *False_iid* case (independence assumption when the errors were in fact correlated); solid triangles show the *False_cid* case (assumption of correlated errors when they were actually independent); and open squares stand for the *True_cid* case (correlated assumption with correlated errors). The symbols get larger with increasing numbers of stations in the

random subnetworks. **Table 6** summarizes the actual coverages obtained from the 10, 20, 30 and 40-station subnetworks and the various cases of locator assumptions.

Table 6. Coverage versus Monte Carlo inputs using various random sub-networks (10, 20, 40, and 40 stations), and locator assumptions.

Locator assumption	Gaussian errors				GEV errors			
	10	20	30	40	10	20	30	40
<i>True_iid</i>	87%	90%	87%	88%	81%	83%	83%	81%
<i>False_iid</i>	78%	61%	46%	40%	71%	51%	37%	32%
<i>False_cid</i>	92%	98%	99%	99%	88%	96%	97%	99%
<i>True_cid</i>	90%	91%	88%	88%	84%	83%	83%	85%

When the locator makes the right assumptions (*True_iid* and *True_cid*) the coverage statistic does not depend on the number of stations and we get about 90% actual coverage (**Figure 35a**); i.e. given the size of the Monte Carlo populations, the coverages are not statistically significantly different than the expected 90%. When the residuals are drawn from the non-Gaussian GEV distribution (**Figure 35b**) the actual coverage decreases by 5-10%. Since the networks are relatively well-balanced with occasional “clumps” of stations, the 10-station sub-networks are close to being independent. However, as the number of stations increases, the correlation structure gets stronger. This is reflected in the rapidly deteriorating coverage statistics when the locator wrongly assumes that the observations are independent (*False_iid*). Thus, there is a severe penalty for ignoring the correlated error structure, which clearly overpowers the deteriorating effect of non-Gaussian errors. On the other hand, the *False_cid* case, when the locator falsely assumes that the errors are correlated when they are in fact independent, is relatively benevolent as it errs on the conservative side. Since the locator thinks that there are fewer independent observations than they actually are, it overestimates the location uncertainties, thus yielding overly conservative coverage statistics.

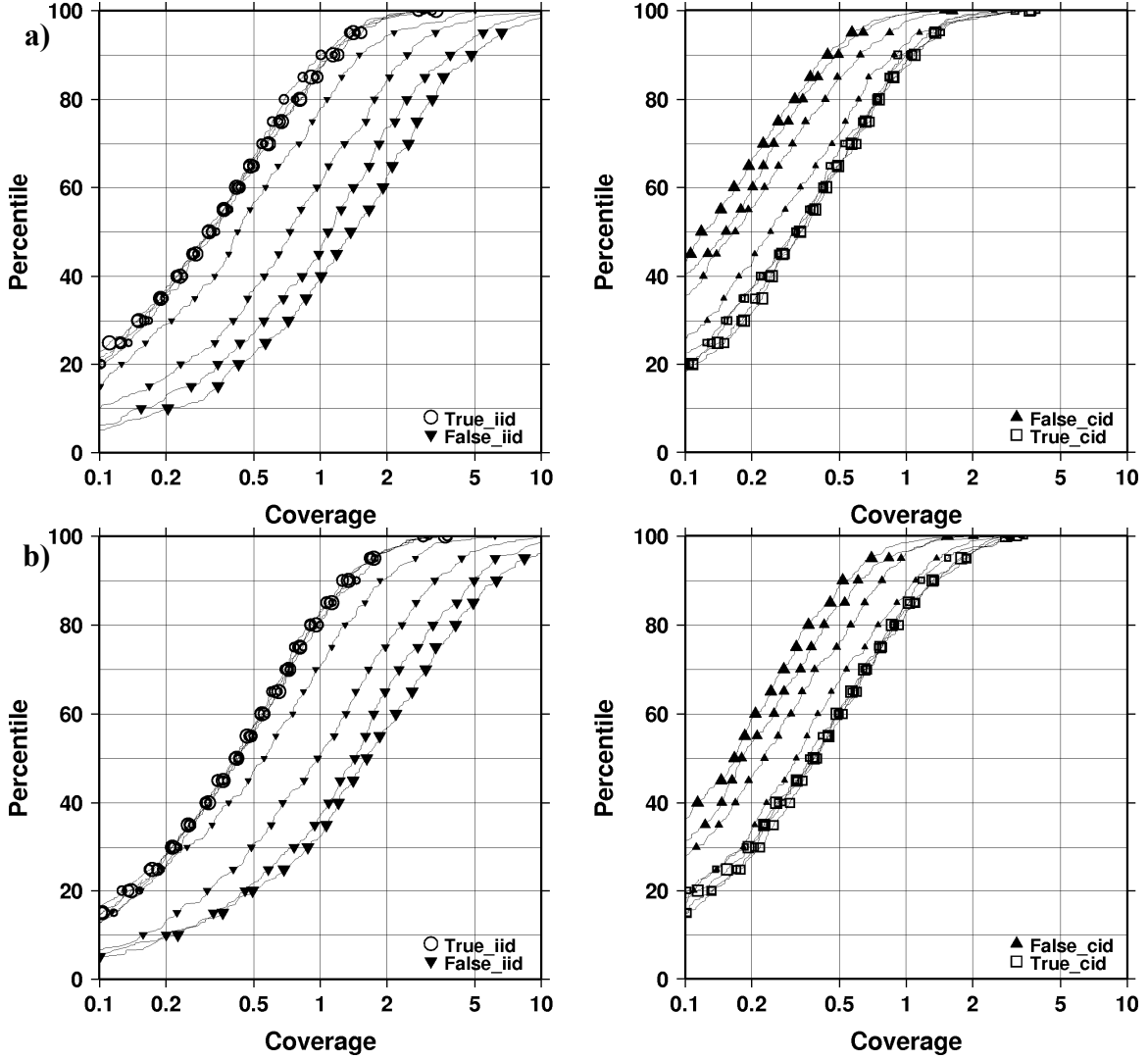


Figure 35. Cumulative coverage parameters from synthetic Monte Carlo location experiments with a) Gaussian, and b) GEV errors. The symbols get increasingly larger for the 10, 20, 30 and 40-station random sub-networks.

4.4. Correlated model errors with mb-based measurement errors and phase pick delay correction

In this last validation test we present relocation results for GT0-2 nuclear explosions and GT5 earthquakes when the full data covariance matrix is constructed from the network covariance matrix to account for correlated model errors, measurement errors are modeled as mb-distance dependent estimates (independent) and phase pick delay corrections are applied. Relocations were conducted with arrival data drawn from the validation data set described above. The results are compared with those from the baseline location algorithm (ignoring correlated model error structure and assuming 1 second measurement errors with no phase pick delay corrections) using uncalibrated (iasp91), and calibrated (CUB2, J362D28) travel-time predictions.

4.4.1. Nuclear explosions

Figure 36 shows the cumulative distributions of mislocation and coverage parameter when regional Pn (**Figure 36a**), teleseismic P (**Figure 36b**) and both Pn and P (**Figure 36c**) were used in the relocations. Note that representation at regional distance range is dominated by NTS explosions so the statistics are somewhat biased towards the western US. Other test sites typically suffer from poor regional station coverage, which explains the sudden increase in mislocations for the upper 20% quantile of events in **Figure 36a**.

In general, calibrated travel-times are responsible for the location improvements, but because of their somewhat optimistic model error estimates they tend to underestimate the formal uncertainties, which leads to deterioration in their coverage statistics. Location improvements due to phase pick delay corrections and the correlated-error model are typically small and secondary to improvements made by the calibrated travel times. Nevertheless, the incorporation of the full data covariance matrix in the location algorithm results in vastly improved formal location uncertainty estimates, and the distribution of the coverage parameter approaches the expected theoretical χ^2 distribution.

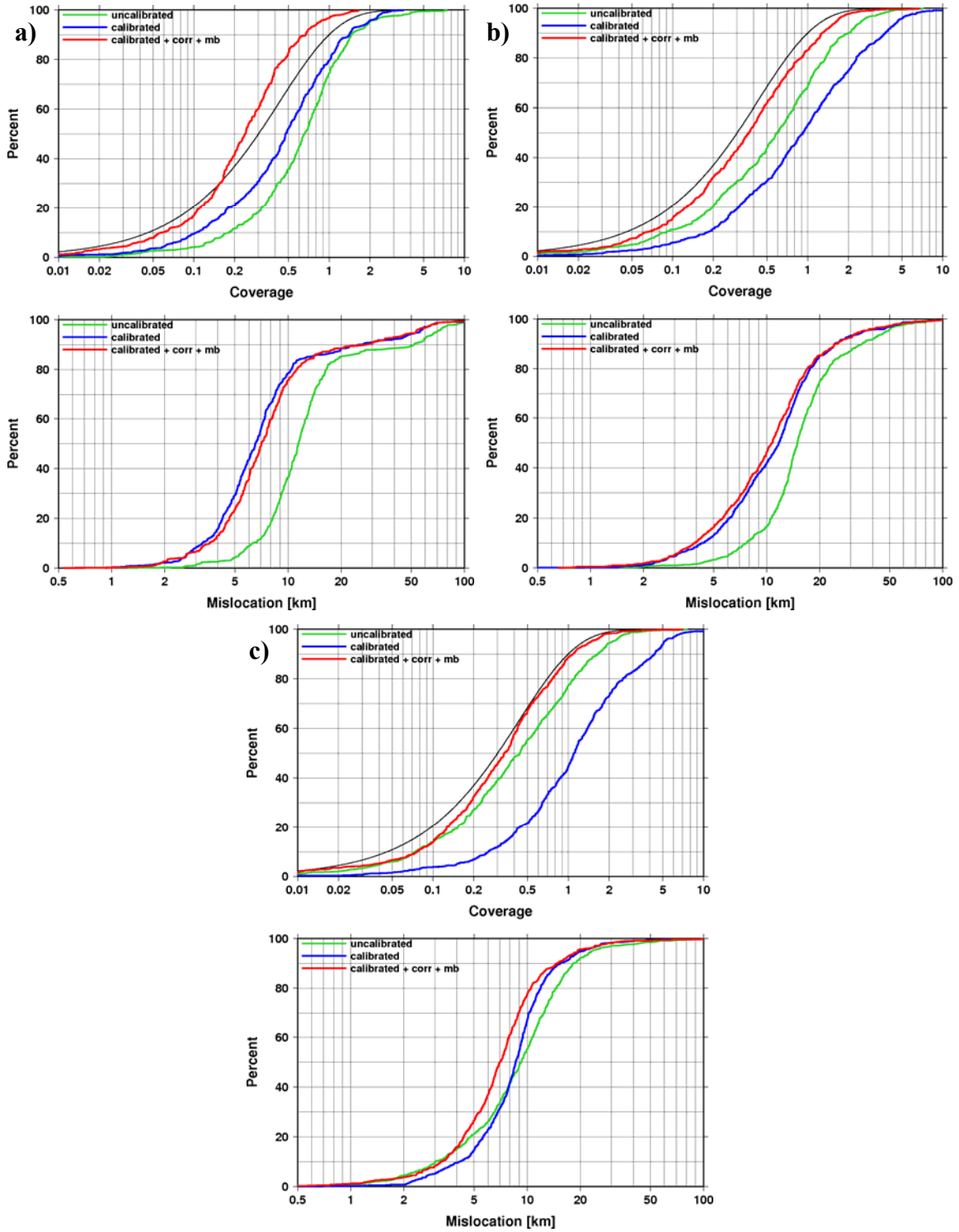


Figure 36. Cumulative distributions of mislocation and coverage for GT0-2 nuclear explosions relocated with a) Pn phases between 0°-20°, b) P between 28°-90°, and c) both Pn and P. Green lines denote results with uncalibrated (iasp91) travel-time predictions assuming independence and 1 second reading errors; blue lines show results with calibrated (CUB2, J362D28) travel-times; red lines represent the relocation results with calibrated travel-times when correlated model errors are accounted for and using mb-based measurement error and bias estimates. Black lines represent the cumulative distribution of the coverage parameter for the expected theoretical χ^2 distribution.

As we noted before, mb-based phase pick delay corrections tend to tighten event clusters by moving small events closer to large ones. This is illustrated with Yucca Flat GT0 explosions in **Figure 37 a-c**, using Pn phases only. The area of the 90% ellipse encompassing the point cloud decreases from 231 km² to 153 km² due to calibrated travel-times; SSSCs with mb-based delay corrections and accounting for correlated model error structure further decreases the ellipse area to 131 km². When teleseismic P readings are also included in the location (**Figure 37 d-f**) the reduction of the “rainbow effect” (the systematic separation between small (red) and large (blue) events) due to the phase pick delay corrections is less pronounced because the amplitude of delay corrections diminishes at teleseismic distances (see **Figure 11**). Nevertheless, taking into account the correlated model error structure reduces the remaining location bias due to the calibrated travel-times.

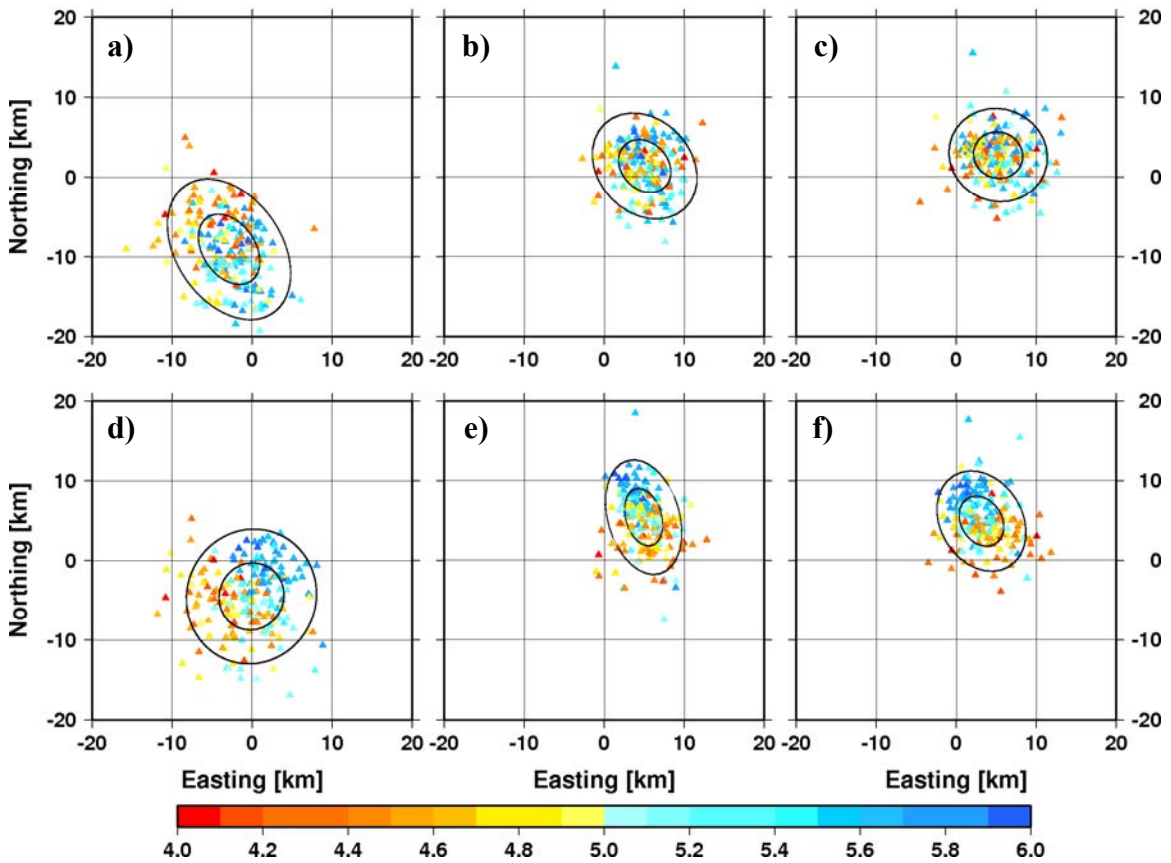


Figure 37. Mislocations of Yucca Flat explosions using a) uncalibrated Pn, b) calibrated Pn travel-time predictions, and c) calibrated Pn travel-times accounting for correlated model errors and mb-based delay and variance estimates. d-f are the same as a-c but using both regional Pn and teleseismic P phases. Events are color-coded by their mb and locations are plotted relative to the GT0 locations in Easting-Northing coordinates. Error ellipses at the one and two sigma levels encompassing the point clouds are also shown.

4.4.2. Earthquakes

Figure 38 shows the cumulative distributions of mislocation and coverage parameter when regional Pn (**Figure 38a**), teleseismic P (**Figure 38b**) and both Pn and P (**Figure 38c**) were used in the relocations of GT5 earthquakes. The dataset of more than 2,000 GT5 earthquakes represents a wide variety of sparse and dense, balanced and unbalanced networks.

As with the explosions, calibrated travel-times are again responsible for the bulk of earthquake location improvements, but they underestimate the formal uncertainties. The expected location improvements due to phase pick delay corrections and accounting for correlated model errors are typically small (1-2 km) and well within the GT5 accuracy. Therefore it is not surprising that they provide no measurable location improvements. Furthermore, because the random picking errors are typically larger than those made on explosions, they weaken the correlation structure (c.f. **Figure 16**). Thus, the improvements in earthquake coverage statistics are less dramatic than for the explosions. Nevertheless, accounting for the correlated model errors provides nearly 90% actual coverage, and the distribution of the coverage parameter closely follows the expected χ^2 distribution.

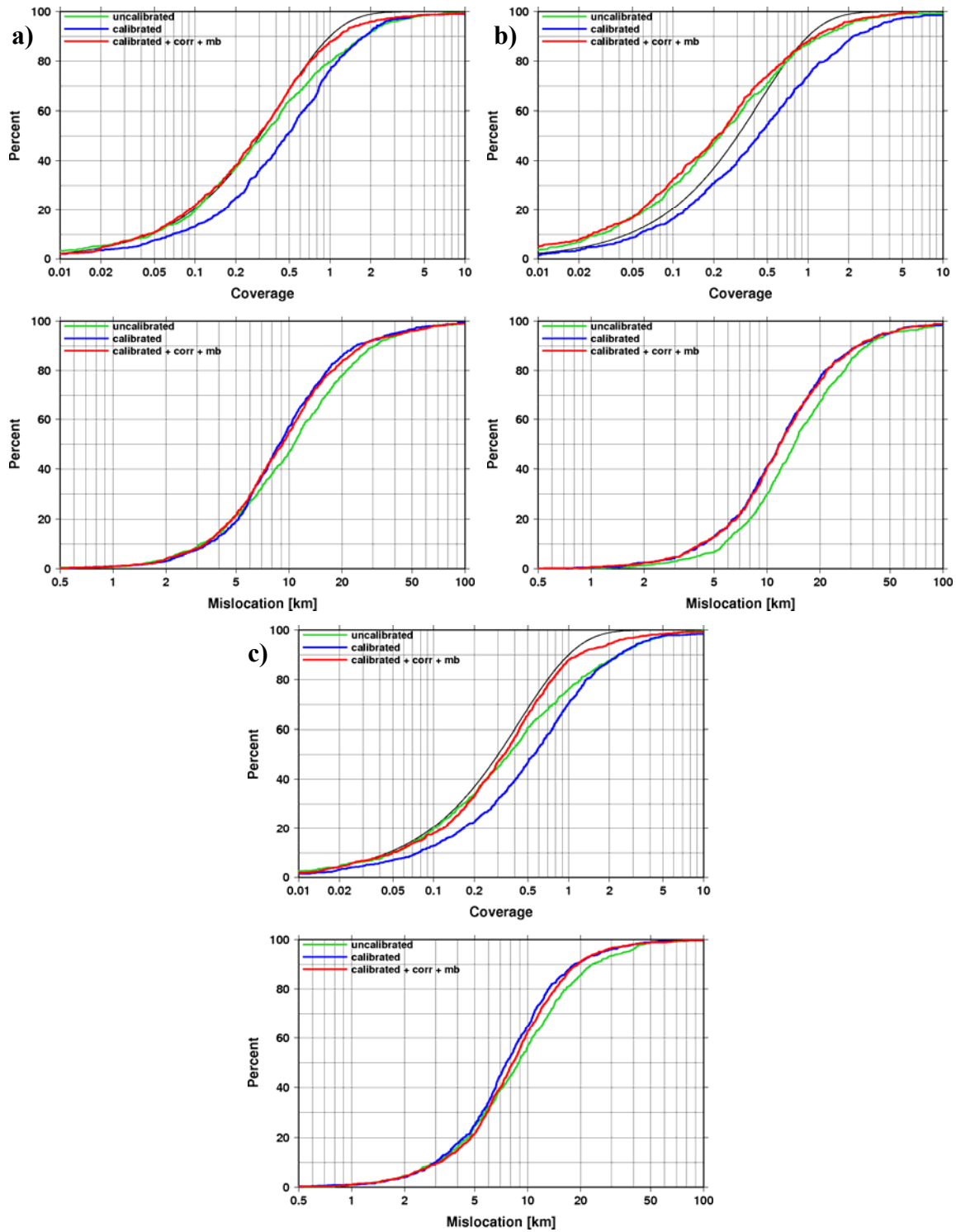


Figure 38. Cumulative distributions of mislocation and coverage for GT5 earthquakes relocated with a) Pn phases between 0°-20°, b) P between 28°-90°, and c) both Pn and P. Green lines denote results with uncalibrated (iasp91) travel-time predictions assuming independence and 1 second reading errors; blue lines show results with calibrated (CUB2, J362D28) travel-times; red lines represent the relocation results with calibrated travel-times when correlated model errors are accounted for and using mb-based measurement error and bias estimates. Black lines represent the cumulative distribution of the coverage parameter for the expected χ^2 distribution with 2 degrees of freedom.

5. DISCUSSION AND CONCLUSIONS

We presented a methodology accounting for the correlated model error structure in a linearized iterative least squares location algorithm with demonstrated improvements in both location and location uncertainty estimates. However, we know that in most methodologies the devil is in the assumptions. Let us reiterate the assumptions we have made and their implications.

- 1) We assumed that the similarity between two ray paths can be approximated by the distance between the two receivers, which results in a stationary correlation structure that does not depend on epicentral distance. The stationarity assumption means that the full data covariance matrix and its inverse needs to be calculated only once. This assumption breaks down at epicentral distances corresponding to discontinuities in the ray parameter where pairs of ray paths rapidly decorrelate (Rodi and Myers, 2007) or sample very different structures. However, as we have shown above, assuming correlation when there is none simply yields more conservative uncertainty estimates.
- 2) We assumed that an isotropic variogram model adequately characterizes the correlation structure. Unless one uses three-dimensional velocity models that explain at least the bulk of 3D variations in the velocity structure in the Earth, this assumption becomes increasingly indefensible for shorter ray paths where there is less averaging over various scale lengths of velocity perturbations. Hence, we used the CUB2 and J326D28 global 3D models to generate Pn and P travel-time predictions. As new, improved generations of 3D models on global, regional and local scales become available, validity of the isotropic assumption will increase.

Having said this, for the long, teleseismic ray paths one may achieve coverage improvements with correlation models without calibrated travel-times. To test this hypothesis, we relocated 640 GT0-2 underground nuclear explosions (stars in **Figure 4**) from various test sites as well as “peaceful” nuclear explosions from the former Soviet Union and the USA. **Figure 39** shows the cumulative distributions of mislocation and coverage using calibrated (J362D28, red lines) and uncalibrated (iasp91, orange lines) travel-time predictions when correlated model errors are accounted for. The green lines show the baseline location results (iasp91 predictions, correlation structure ignored), and blue lines show the cumulatives using calibrated travel-time predictions but still assuming independent errors. The correlated cumulatives are nearly always better than their respective independent cumulatives (uncalibrated vs calibrated).

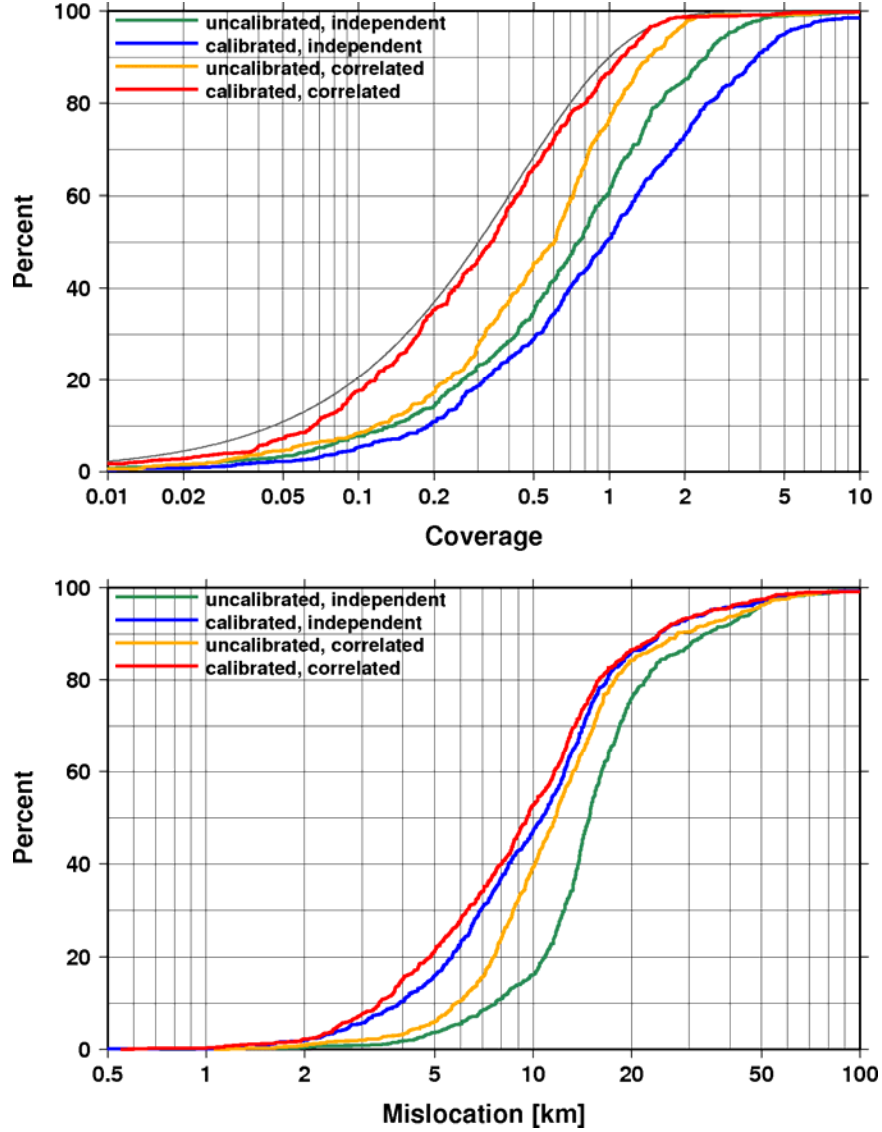


Figure 39. Mislocation and coverage for 640 GT0-2 nuclear explosions using teleseismic (28° - 90°) P phases. The thin grey line represents the cumulative distribution of the coverage parameter for the theoretical χ^2 distribution with 2 degrees of freedom. Green lines denote the baseline results with uncalibrated (iasp91) travel-time predictions when correlated error structure is ignored. Orange lines represent the results with uncalibrated travel-time predictions when correlated error structure is accounted for. Blue lines are results using calibrated (J362D28) travel-time predictions but assuming independent errors. Red lines are cumulatives when correlated errors are accounted for using calibrated travel-time predictions. Coverage and location are improved when correlated structures are accounted for regardless of calibrated or uncalibrated travel times.

Table 7 gives a summary of median mislocations and actual coverage statistics for the four different cases. Even when uncalibrated travel-time predictions are used, accounting for the correlation structure yields significant improvements in coverage (from 60% to 75%) and consistent improvements in location. Calibrated travel-time predictions are responsible for the bulk of the location improvements. However, they marginally improve coverage statistics. One could scale the error

ellipses so that 90% coverage is achieved, but because the shape of the empirical coverage cumulative differs from that expected for a χ^2 distribution with 2 degrees of freedom (thin line), this would only result in erroneous coverage estimates at other significance levels. As **Figure 39** shows, it is the representation of the correlated model error structure that leads to significant improvements in both actual coverage at the 90th percentile level and the correct shape of the cumulative distribution, no matter whether calibrated or uncalibrated travel-time predictions are used. Hence, to achieve the optimal results, one should both use calibrated travel-times and account for correlated model errors. In fact, using 3-D model travel-time predictions and taking into account the correlated error structure we obtain the largest location improvements and almost hit the mark by achieving 87% actual coverage, with the cumulative distribution curve closely following the expected theoretical shape (given the 640 event population 87% is not statistically significantly different than the expected 90%).

Table 7. Median mislocation and actual coverage using combinations of travel-time predictions and location algorithm assumptions.

Locator Assumptions	Median mislocation (km)		Actual coverage	
	<i>Independent Assumption</i>	<i>Correlated Assumption</i>	<i>Independent Assumption</i>	<i>Correlated Assumption</i>
<i>Uncalibrated Travel-times</i>	14.9	11.7	60%	76%
<i>Calibrated Travel-times</i>	10.7	9.6	51%	87%

Another way to present the relocation results is to quote the percentage of event locations that were improved or deteriorated significantly with respect to the baseline algorithm (uncalibrated travel-times, assumption of independent errors). Out of 640 explosions 57% locations are improved by more than 2 km (more than GT0-2 accuracy) while only 11% are deteriorated by more than 2 km when using uncalibrated travel-time predictions but accounting for the correlation structure; 69% and 10% of the locations were improved/deteriorated by more than 2 km using calibrated travel-times and assuming independent errors. When we used calibrated travel-times and accounted for the correlated errors, 72% and 9% of the locations were improved/deteriorated by more than 2 km. If we set the comparison threshold to 5 km, accounting for the correlated structure alone improved 32% and deteriorated only 3% of the locations by more than 5 km. Calibrated travel-time predictions alone improved a significantly larger number of locations (53%) than deteriorated (4%). Finally, using calibrated travel-times and accounting for correlated errors improved 55% and deteriorated 3% of the locations by more than 5 km, respectively.

- 3) Throughout this paper we utilized a location algorithm that assumed zero-mean, Gaussian picking errors. Many researchers have noted that the distribution of

picking errors is non-Gaussian. We have shown that the heavy tailed Generalized Extreme Value distribution provides a better fit to the GT residual distribution. The skewed shape of the residual distribution is attributed to the fact that signals are typically picked late with decreasing event size, or more precisely, with decreasing signal-to-noise (SNR) ratio (Anderson, 1982; Douglas et al., 1997, 2005; Kvaerna, 1996). We demonstrated that accounting for the systematic pick delay (using either mb-distance or SNR-based measurement error models) eliminates bias in the residual distribution, i.e. the delay corrections make the residual distribution zero-mean. The estimates of measurement error bias and variance also make the distribution more Gaussian-like, thus they facilitate the location algorithm zero-mean Gaussian error assumption.

- 4) We assumed that the correlated model errors may be described by a non-diagonal covariance matrix. This implicitly implies Gaussian errors and a linear correlation structure which does not allow for tail dependence. Tail dependence may occur when outliers act in unison to spoil the location. With GT data, we have shown that the effect of non-Gaussian marginal distributions, coupled via a linear correlation structure, becomes significant only at the highest percentile levels. We have demonstrated that the bulk of the model error structure can be characterized by the correlation matrix. Since copula methodology allows for non-linear dependence structures between arbitrary marginal distributions, further refinements concerning the dependence structure are possible using the correlated error iterative least-squares algorithm coupled to a Monte Carlo algorithm to exhaustively contour significance levels.
- 5) We have also postulated that generic variogram models can be used to construct network covariance matrices that will perform reasonably well anywhere in the globe. We have shown that this approach indeed yields significant improvements in location uncertainty estimates as well as location improvements for unbalanced networks. Of course, one can always do better by calibrating a specific region. However, for organizations working in an operational environment where large numbers of events are located near real-time, the application of our simplified approach may result in improved event bulletins.

We have shown that ignoring the correlated error structure due to unmodeled velocity lateral heterogeneities results in unreliable location uncertainty estimates, and for unbalanced networks, location bias. Unfortunately, most location algorithms (linearized iterative, non-linear grid-search) and their error estimators (linearized Gaussian or non-Gaussian Monte Carlo) do exactly this.

We demonstrated a location algorithm that takes into account the correlated model error structure described by a non-diagonal covariance matrix. The algorithm provides more accurate formal uncertainty estimates and small, but consistent location improvements for both sparse and dense networks. Our synthetic Monte Carlo experiments have shown that coverage statistics rapidly deteriorate with increasing number of correlated stations when the correlated error structure is ignored. On the other

hand, assuming a correlated model error structure when there is none errs on the conservative side by overestimating the formal uncertainties.

We have shown that the effects of SNR-dependent picking delay and variance estimates derived from either direct estimates of SNR (sta/lta) or SNR surrogates (A/T or mb) on event locations is less than those improvements achieved by calibrated travel-times and accounting for correlated model errors; incorporation of SNR dependent error models do tighten event clusters by reducing the systematic location biases between small and large events but do not significantly reduce total cluster bias. A direct test of the hypothesis that incorporation of sta/lta based SNR dependent errors will reduce location biases must await a more extensive GT event and arrival data set.

While non-Gaussian errors consistently deteriorate both location and coverage, their neglect is secondary relative to the penalty paid for ignoring the correlated error structure. This, however, can by no means be interpreted as implying that the non-Gaussian nature of error distributions is negligible. Further improvements in location uncertainty estimates and reductions in location bias should be achieved by modeling the measurement errors as non-Gaussian, skewed distributions, and perhaps accounting for non-linear model error dependence structures.

ACKNOWLEDGEMENTS

The ground truth earthquake clusters used in this study were developed under the US Air Force Research Laboratory contract FA8718-04-C-0020. We thank Bob Engdahl, Mike Ritzwoller (Univ. Colorado at Boulder) and Mike Antolik (Quantum Technology Services, Inc.) for making available the EHB bulletin, the CUB2 and J362D28 global Earth models. We used the open source R package *copula* by Yan (2007), and generated figures using the Generic Mapping Tool (GMT4.2.1) software developed by Wessel and Smith (1991).

REFERENCES

- Anderson, K.R. (1982). Robust earthquake location using M estimates, *Phys. Earth Planet. Int.*, **30**, 119-130.
- Antolik, M., Y.J. Gu, G. Ekström and A. Dziewonski (2003). J362D28: A new joint model of compressional and shear velocities in the Earth's mantle, *Geophys. J. Int.*, **153**, 443-466.
- Bahavar, M., B.W. Barker, T.J. Bennett, J.R. Bowman, H. Israelsson, B. Kohl, Y-L. Kung, J.R. Murphy, R.G. North, V. Oancea, M. O'Brien and G. Shields (2007). Enhanced resources of the SMDC monitoring research program for source information and data acquisition, *29th Seismic Research Review: Nuclear Explosion Monitoring: Building the Knowledge Base*, Denver, Colorado, 227-236, LA-UR-07-5613.
- Billings, S.D., M.S. Sambridge and B.L.N. Kennett (1994). Error in hypocenter location: Picking, model and magnitude dependence, *Bull. Seism. Soc. Am.*, **84**, 1978-1990.
- Bondár, I., S.C. Myers, E.R. Engdahl and E.A. Bergman (2004a). Epicenter accuracy based on seismic network criteria, *Geophys. J. Int.*, **156**, 483-496, doi 10.1111/j.1365-246X.2004.02070.x.
- Bondár, I., E.R. Engdahl, X. Yang, H.A.A. Ghalib, A. Hofstetter, V. Kirichenko, R. Wagner, I. Gupta, G. Ekström, E. Bergman, H. Israelsson and K. McLaughlin (2004b). Collection of a reference event set for regional and teleseismic location calibration, *Bull. Seism. Soc. Am.*, **94**, 1528-1545.
- Bondár, I., E. Bergman, B. Kohl, Y-L. Kung, H. Israelsson, K. McLaughlin and E.R. Engdahl (2007). Global ground truth data set with waveform and improved arrival data, *Science Applications International Corp.*, Final Report, AFRL-RV-HA-TR-2007-1101.
- Bondár, I., E. Bergman, E.R. Engdahl, B. Kohl, Y-L. Kung and K. McLaughlin (2008). A hybrid multiple event location technique to obtain ground truth event locations, submitted to *Geophys. J. Int.*
- Buland, R. (1986). Uniform reduction analysis, *Bull. Seism. Soc. Am.*, **76**, 217-230.
- Chang, A.C., R.H. Shumway, R.R. Blandford and B.W. Barker (1983). Two methods to improve location estimates – preliminary results, *Bull. Seism. Soc. Am.*, **73**, 281-295.
- Cleveland, W. S. (1979), Robust locally weighted regression and smoothing scatterplots, *J. Amer. Statist. Assoc.*, **74**, 829-836.
- Douglas, A., D. Bowers and J.B. Young (1997). On the onset of P seismograms, *Geophys. J. Int.*, **129**, 681-690, 1997.
- Douglas, A., J. B. Young, D. Bowers, and P. D. Marshall (2005a). An Analysis of P Travel Times for Nevada Test Site Explosions Recorded at Regional Distances, *Bull. Seism. Soc. Am.*, **95**, 941-950.
- Douglas, A., J.B. Young, D. Bowers and M. Lewis (2005b). Variation in reading error in P times for explosions with body-wave magnitude, *Phys. Earth Planet. Int.*, **152**, 1-6.

- Douglas, A., A.M. O'Mongain, D. Porter, and J.B. Young (2005c). An analysis of P times reported in the Reviewed Event Bulletin for Chinese underground explosion, *Geophys. J. Int.*, **163**, 559-566.
- Engdahl, E.R., R.D. van der Hilst and R.P. Buland (1998). Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seism. Soc. Am.*, **88**, 722-743.
- Evernden, J. (1969). Precision of epicenters obtained by small numbers of world-wide stations, *Bull. Seism. Soc. Am.*, **59**, 1365-1398.
- Fisk, M. (2002). Accurate locations of nuclear explosions at the Lop Nor test site using alignment of seismograms and IKONOS satellite imagery, *Bull. Seism. Soc. Am.*, **92**, 2911-2925.
- Flanagan, M.P., S.C. Myers and K.D. Koper (2007). Regional travel-time uncertainty and seismic location improvement using a three-dimensional a priori velocity model, *Bull. Seism. Soc. Am.*, **97**, 804-825.
- Flinn, E. (1965). Confidence regions and error determinations for seismic event location, *Rev. Geophys.*, **3**, 157-185.
- Freedman, H. W. (1966). A statistical discussion of P_n residuals from explosions, *Bull. Seism. Soc. Am.*, **56**, 677-695.
- Frees, E.W. and E.A. Valdez (1998). Understanding relationships using copulas, *North Am. Actuarial J.*, **2**, 1-25, 1998.
- Genest, C., L-P. Rivest (1993). Statistical inference procedures for bivariate Archimedean copulas, *J. Am. Stat. Assoc.*, **88**, 1034-1043.
- Herrin, E. and J. Taggart (1968). Source bias in epicenter determination, *Bull. Seism. Soc. Am.*, **58**, 1791-1796.
- International Seismological Centre, On-line Bulletin, <http://www.isc.ac.uk/Bull>, Internatl. Seism. Cent., Thatcham, United Kingdom, 2001.
- Jenkinson, A.F. (1955). The frequency distribution of the annual maximum (or minimum) of meteorological elements. *Quart. J. R. Met. Soc.*, **81**, 158-171.
- Joe, H. (1997). Multivariate Models and Dependence Concepts, Chapman & Hall, London.
- Jordan, T.H. and K.A. Sverdrup (1981). Teleseismic location techniques and their application to earthquake clusters in the South-central Pacific, *Bull. Seism. Soc. Am.*, **71**, 1105-1130.
- Kennett, B. and E.R. Engdahl (1991). Travel Times for Global Earthquake Location and Phase Identification, *Geophys. J. Int.*, **105**, 429-465.
- Kohl, B., H. Israelsson and M. Fisk (2003). Advanced concepts for nuclear explosion monitoring applied to regional events in the Arctic region, 25th *Seismic Research Review: Nuclear Explosion Monitoring: Building the Knowledge Base*, Tucson, Arizona, 227-236, LA-UR-03-6029.

- Kværna, T. (1996). Time shifts of phase onsets caused by SNR variations, *NORSAR Sci. Rep.* 2-95/96, 143-152.
- Lilwall, R. C. and J. M. Neary (1986). Redetermination of Earthquake Body-Wave Magnitudes (mb) using ISC Bulletin Data. AWRE Report No. 0 21/85, HMSO, London.
- Lomnitz, C. (1995). Comments on “Errors in hypocenter location: Picking, model, and magnitude dependence,” by S.D. Billings, M.S. Sambridge, and B.L.N. Kennett, *Bull. Seism. Soc. Am.*, **85**, 1527-1528.
- Morozov, I.B., E.A. Morozova, S.B. Smithson, P.G. Richards, V.I. Khalturin and L.N. Solodilov (2005). 3D first-arrival regional calibration model of Northern Eurasia, *Bull. Seism. Soc. Am.*, **95**, 951-964.
- Murphy, J.R., W. Rodi, M. Johnson, D.D. Sultanov, T.J. Bennett, M.N. Toksöz, V. Ovtchinnikov, B.W. Barker, D.T. Retier, A.C. Rosca and Y. Shchukin (2005). Calibration of International Monitoring System (IMS) stations in Eastern Asia for improved seismic event location, *Bull. Seism. Soc. Am.*, **95**, 1535-1560.
- Myers, S.C. and C.A. Schultz (2000). Improving sparse network seismic location with Bayesian kriging and teleseismically constrained calibration events, *Bull. Seism. Soc. Am.*, **90**, 199-211.
- Nelsen, R.B. (1999). An introduction to copulas, *Lecture Notes in Statistics*, **139**, Springer-Verlag, New York.
- Pavlis, G.L. (1986). Appraising earthquake hypocenter location errors: a complete, practical approach for single-event locations, *Bull. Seism. Soc. Am.*, **76**, 1699-1717.
- Reiter, D., W. Rodi and M. Johnson (2005). Development of a tomographic upper-mantle velocity model beneath Pakistan and Northern India, *Bull. Seism. Soc. Am.*, **95**, 926-940.
- Ritzwoller, M.H., N.M. Shapiro, E.A. Levshin, E.A. Bergman and E.R. Engdahl (2003). Ability of a global three-dimensional model to locate regional events, *J. Geophys. Res.*, **108**(B7), 2353, DOI 10.1029/2002JB002167.
- Rodi W.L. and S.C. Myers (2007). Modeling travel-time correlations based on seismicity kernels and correlated velocity anomalies, *29th Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies*, Denver, Colorado, 463-471, LA-UR-07-5613.
- Röhm, A.H.E., J. Trampert, Paulssen, H., and R. Sneider (1999). Bias in reported seismic arrival times deduced from the ISC Bulletin, *Geophys. J. Int.*, **137**, 163-174.
- Salvadori, G., C. De Michele, N.T. Kottegoda and R. Rosso (2007). Extremes in Nature: An Approach Using Copulas, *Water Science and Technology Library*, **56**, Springer.
- Shapiro, N.M. and M.H. Ritzwoller (2004). Thermodynamic constraints on seismic inversions, *Geophys. J. Int.*, **157**, 1175-1188, doi 10.1111/j.1365-246X.2004.02254.x.

- Sklar, A. (1959). Fonctions de repartition à n dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris*, **8**, 229-231.
- Song, P.X.K. (2000). Multivariate dispersion models generated from Gaussian copulas, *Scand. J. Statist.*, **27**, 305-320.
- Springer D. L., Pawloski G. A., Ricca J. L., Rohrer R. F., and D. K. Smith, (2002). Seismic Source Summary for All U.S. Below-Surface Nuclear Explosions, *Bull. Seism. Soc. Am.*, **92**, 1806 – 1840.
- Veith, K F and G E Clawson, (1972). Magnitude from short-period P-wave data, *Bull. Seism. Soc. Am.*, **62**, 435 - 452.
- Venables, W. N. and B. D. Ripley (2002). Modern Applied Statistics with S, Fourth edition, Springer.
- Wessel, P. and W. H. F. Smith (1991). Free software helps map and display data, EOS Trans. AGU, 72, 441.
- Yan, J. (2007). Enjoy the joy of copulas: With a package copula, *J. Stat. Soft*, **21**, 4, 1-21.
- Yang, X., I. Bondár, K. McLaughlin and R.G. North (2001). Source Specific Station Corrections for regional phases at Fennoscandian stations, *Pure Appl. Geophys.* **158**, 35-57.
- Yang, X., I. Bondár, J. Bhattacharyya, M. Ritzwoller, N. Shapiro, M. Antolik, G. Ekström, H. Israelsson and K. McLaughlin (2004). Validation of regional and teleseismic travel-time models by relocating GT events, *Bull. Seism. Soc. Am.*, **94**, 897-919.

APPENDIX A

Copulas were used in the preceedings to generate correlated non-Gaussian random numbers, and to construct robust variogram estimates. In the following we outline these two copula applications.

A.1 Generating correlated multivariate non-Gaussian random numbers

In order to generate multivariate random numbers one must construct the joint probability distribution. While this is relatively straightforward for the multivariate normal distribution, it becomes problematic for the general non-Gaussian case. We want to generate multivariate random numbers from a joint distribution with prescribed marginal distributions where the dependence structure is described by a correlation matrix. Copula methodology allows just this.

The general copula theory is summarized in Joe (1997), Nelsen (1999) and Salvadori et al. (2007). A copula is a joint distribution function of standard uniform, $u_k \in [0,1]$, random variables:

$$C(u_1, \dots, u_d) = \Pr\{U_1 \leq u_1, \dots, U_d \leq u_d\} \quad (\text{A.1})$$

Hence, using the probability integral transformation, a copula evaluated at $u_l = F_l(x_l), \dots, u_d = F_d(x_d)$ is identical to the joint distribution function

$$C(F_1(x_1), \dots, F_d(x_d)) = F(x_1, \dots, x_d) \quad (\text{A.2})$$

Sklar's (1959) theorem states that for a d -dimensional distribution function F with margins F_1, \dots, F_d there exists a copula such that

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \quad (\text{A.3})$$

Sklar's theorem therefore implies that copulas separate marginal behavior from their dependence structure. This separation of the dependence structure from the marginal distributions is perhaps more apparent in the form of the joint probability density function:

$$f(x_1, \dots, x_d) = c(F_1(x_1), \dots, F_d(x_d)) \prod_i f_i(x_i) \quad (\text{A.5})$$

where

$$c(u_1, \dots, u_d) = \frac{\partial^d C(u_1, \dots, u_d)}{\partial u_1 \dots \partial u_d} \quad (\text{A.6})$$

denotes the copula density function. Thus, the joint distribution is expressed in terms of its respective marginal distributions and a copula function that binds (couples) them together. A substantial advantage of copulas is that the marginal distributions, $F_k(x_k)$ may come from different families. This construction allows considering marginal distributions and dependence as two separate but related issues. Note that for independent variables the copula is simply the product of the marginal distribution functions. Then the copula density function becomes unity, and the joint probability density function of the independent random variables reduces to the familiar form of the product of the marginal probability density functions.

Thus, generating multivariate random numbers from given marginal distributions and dependence structure is straightforward with copula methodology. By virtue of Sklar's theorem, it is sufficient to generate uniform random variables (u_1, \dots, u_d) whose joint distribution is the copula $C(u_1, \dots, u_d)$. Then, using the quantile functions of the marginal distributions the sample drawn from the joint distribution $F(x_1, \dots, x_d)$ is obtained by

$$(x_1, \dots, x_d) = (F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)) \quad (\text{A.7})$$

The simulation of uniform random variables for a given copula can be done by iterative conditioning which takes advantage of the fact that joint conditional probability distributions are easily expressed by copulas. The algorithm starts with the generation of d independent uniform, $v_k \in [0,1]$, random variates (v_1, \dots, v_d) . Then it recursively generates u_k using the relation $u_k = C_{(u_1, \dots, u_{k-1})}^{-1}(v_k)$ with

$$C_{(u_1, \dots, u_{k-1})}(u_k) = \Pr\{U_k \leq u_k \mid (U_1, \dots, U_{k-1})\} = \frac{\partial_{(u_1, \dots, u_{k-1})}^{k-1} C(u_1, \dots, u_k, 1, \dots, 1)}{\partial_{(u_1, \dots, u_{k-1})}^{k-1} C(u_1, \dots, u_{k-1}, 1, \dots, 1)} \quad (\text{A.8})$$

In our Monte Carlo experiments we used the open source R package *copula* by Yan (2007) to generate multivariate random variables drawn from either normal or Generalized Extreme Value marginal distributions bound together by the Gaussian copula. The Gaussian copula is defined as

$$C(u_1, \dots, u_d, R) = \Phi_d(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d) \mid R) \quad (\text{A.9})$$

where Φ_d is the standard multivariate normal distribution, Φ^{-1} is the quantile function of the standard normal distribution, and R is the correlation matrix. The joint probability density function is then written (Song, 2000) as

$$f(x_1, \dots, x_d) = |R|^{-1/2} \exp\left\{\frac{1}{2} q^T (I_d - R^{-1}) q\right\} \prod_i f_i(x_i) \quad (\text{A.10})$$

where $q = (q_1, \dots, q_d)^T$ with $q_i = \Phi^{-1}(u_i) = \Phi^{-1}(F_i(x_i))$.

A.2 Robust variogram model estimation

We define the robust variogram model as the median regression curve of the residual difference squares for station pairs of common events with respect to station separation: $\gamma(\Delta(sta_i, sta_j)) = \text{median}((\delta t_i - \delta t_j)^2)$. We also require that the variogram model be continuous and monotonically increasing. The copula formalism offers an elegant way to construct joint conditional probability distributions and derive quantile regression curves of y subject to x (Frees and Valdez, 1998). In our case $x = \Delta(sta_i, sta_j)$, $y = (\delta t_i - \delta t_j)^2$ and $u = F_1(x)$, $v = F_2(y)$. The p^{th} quantile regression curve is defined as

$$y_p = F_2^{-1}(v_p) \quad (\text{A.11})$$

where v_p is the solution of the equation

$$C(v_p \mid u) = \int_0^{v_p} \frac{\partial^2 C(u, v)}{\partial u \partial v} dv = \frac{\partial C(u, v_p)}{\partial u} = p \quad (\text{A.12})$$

Setting p to 0.5 yields the median regression curve of y subject to x .

In order to obtain a variogram model we first determined which bivariate copula of the widely used Archimedean family explains the dependence structure between station separation and squared residual differences. Copulas of the form

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)) \quad (\text{A.13})$$

are called Archimedean copulas, where φ is a convex, decreasing function with domain $(0, 1]$ and range $[0, \infty)$ such that $\varphi(1) = 0$. The function φ is called the generator function, which uniquely determines an Archimedean copula. φ is a function of the copula parameter, α . The copula parameter characterizes the dependence structure and in many cases it is directly related to scale-invariant measures of association, such as Spearman's ρ (rank correlation) and Kendall's τ (the measure of relative difference between pairs of concordant and discordant random variables). The main attraction of the Archimedean representation is that it reduces the study of a multivariate copula to a single univariate function, φ , and several families of Archimedean copulas are generally used (for a more complete discussion of Archimedean copulas see Nelsen (1999)). Since φ is a function of the copula parameter, α , identifying φ is equivalent to identifying the Archimedean copula itself.

Genest and Rivest (1993) described a procedure to identify the form of φ from a sample of bivariate observations. The procedure is based on the comparison of the one-dimensional distribution function, often called Kendall's measure or Kendall's distribution, $K_\alpha(t) = \Pr\{C(u, v | \alpha) < t\}$ with its non-parametric estimate. For an Archimedean copula, this distribution function is related to the generator function through the expression

$$K_\alpha(t) = t - \varphi_\alpha(t) / \varphi'_\alpha(t) \quad (\text{A.14})$$

The parametric estimate of $K_\alpha(t)$ is constructed by estimating Kendall's τ from the sample, which is used to get an estimate for the copula parameter α , which yields φ_α and K_α for each family of parametric copulas. The non-parametric estimation K_n is given by

$$K_n(t) = \frac{\sum_i 1(\omega_i \leq t)}{n}, \quad 0 < t < 1 \quad (\text{A.15})$$

where

$$\omega_i = \frac{\sum_j 1(x_j < x_i, y_j < y_i)}{n-1}, \quad i = 1, \dots, n \quad (\text{A.16})$$

The best copula is the one that minimizes the distance between K_α and K_n . We found that the Clayton copula provides the best fit to the data at both regional and teleseismic distance ranges. The Clayton copula is defined as

$$C(u, v) = \max\left(\left[u^{-\alpha} + v^{-\alpha} - 1\right]^{-1/\alpha}, 0\right) \quad (\text{A.17})$$

with generator function

$$\varphi(t) = \frac{(t^{-\alpha} - 1)}{\alpha} \quad (\text{A.18})$$

The relationship between the Clayton copula parameter, α ($\alpha > 0$), and Kendall's τ is given by

$$\tau = \frac{\alpha}{\alpha + 2}, \text{ or } \alpha = \frac{2\tau}{1 - \tau} \quad (\text{A.19})$$

Note that the Clayton copula allows only for positive dependence, with $\alpha \rightarrow 0$ representing independence and $\alpha \rightarrow \infty$ full linear correlation. The copula parameter, α , varies between 0.1 and 0.35 for the earthquake and nuclear explosion clusters for which we derived individual Pn and P variogram models. Our generic Pn and P variogram models are characterized by α values of 0.27 and 0.12, respectively.

For the Clayton copula, the p^{th} quantile regression curve is expressed in a closed form:

$$C(v_p | u) = u^{-\alpha-1} (u^{-\alpha} + v_p^{-\alpha} - 1)^{-(\alpha+1)/\alpha} = p \quad (\text{A.20})$$

Therefore the median regression curve ($p = 0.5$) becomes

$$v_{0.5} = \left(1 + u^{-\alpha} (0.5^{-\alpha/(\alpha+1)} - 1)\right)^{-1/\alpha} \text{ and } y_{0.5} = F_2^{-1}(v_{0.5}) \quad (\text{A.21})$$

The Clayton copula exhibits lower tail dependence, conveniently describing the fact that with decreasing station separation the residual differences become increasingly correlated.